

# Open Research Online

---

The Open University's repository of research publications and other research outputs

## Testing and Analysis of a Computational Model of Human Rhythm Perception

### Thesis

#### How to cite:

Angelis, Vassilis (2014). Testing and Analysis of a Computational Model of Human Rhythm Perception. PhD thesis The Open University.

For guidance on citations see [FAQs](#).

© 2014 Vassilis Angelis

Version: Version of Record

Link(s) to article on publisher's website:  
<http://dx.doi.org/doi:10.21954/ou.ro.00009e47>

---

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

---

[oro.open.ac.uk](http://oro.open.ac.uk)

TESTING AND ANALYSIS OF A  
COMPUTATIONAL MODEL OF HUMAN  
RHYTHM PERCEPTION

VASSILIS ANGELIS

A thesis submitted to  
The Open University, UK  
for the degree of

Doctor of Philosophy

Computing Department  
Faculty of Mathematics, Computing and Technology  
The Open University, UK

Supervisors: Simon Holland, Martin Clayton, Paul J. Upton

August 2013 –



Dedicated to all my family



## ABSTRACT

---

This thesis presents an original methodology, as detailed below, applied to the testing of an existing computational model of human rhythm perception. Since the computational model instantiates neural resonance theory (Large and Snyder, 2009), the thesis also tests that theory. Neural resonance theory is a key target for testing since, as contrasted with many other theories of human rhythm perception, it has relatively strong physiological plausibility. Rather than simply matching overt features of human rhythm perception, neural resonance theory shows how these features might plausibly emerge from low-level properties of interacting neurons.

The thesis tests the theory using several distinct research methods. The model stood up well to each family of tests, subject to limitations that are analysed in detail.

Firstly, the responses of the model to several types of polyrhythmic stimuli were compared with existing empirical data on human responses regarding beat identification to the same stimuli, at a variety of tempi. Polyrhythmic stimuli closely resemble real life complex rhythmical stimuli such as music, and were used for the first time to test the model. It was found that the set of categories of response predicted by the model matched human behaviour.

Secondly, the model was systematically analysed by exploring the degree of dependence of its behaviour on the values of its parameters (sensitivity analysis). The behaviour of the model was found to retain consistency in the face of systematic numerical manipulation of its parameters.

Thirdly, the behaviour of the model was compared to that of related models. In particular, the focal computational model, which balances physiological plausibility with mathematical convenience, was compared with other models that relate more directly to brain physiology. In each case, all relevant behaviours were found to be closely in line.

Lastly, the outputs of the model under polyrhythmic stimuli were analysed to make new testable predictions about previously unobserved human behaviour regarding the time it takes for people to perceive beat in polyrhythms. These predictions led to the design and conduction of new human experimental studies. It was found that the model had successfully predicted previously unobserved aspects of human behaviour, more specifically it predicted the timescale within which people start to perceive beat in a given polyrhythmic stimulus.



## PUBLICATIONS

---

Chapter 6 and parts of Chapter 5 are based on the following journal paper.

Angelis, V., Holland, S., Upton, J. P., and Clayton, M. (2013). Testing a computational model of rhythm perception using polyrhythmic stimuli, In *Journal of New Music Research*, 42(1): 47-60.





*...for being so inspiring  
that you get our neurons firing  
and spontaneously re-wiring  
OU, we owe you*  
— Open University 40th Anniversary Poem  
by Matt Harvey

## ACKNOWLEDGMENTS

---

Big thank you to my supervisors Martin Clayton, Paul J. Upton, and Simon Holland for the sheer support to me from the very beginning to the very end, and for helping me to realise the most out of my doctoral studies. Always indebted.

Thank you to Agota Solyomi for coping and bearing with me all this time and for studying together with me over the weekends. Santiago Ramon y Cajal would have been very pleased to know you.

Thanks to Marian Petre for setting me up in the thesis writing-up orbit, for listening to me and for sharing her insights with all of us. We have been fortunate to have you around. Thanks to fellow PhD students; thank you Andy Milne for keeping up the good work and for bringing the brocots together, thank you Jean Barbier for really caring about *things*, thank you Alexandra Pawlik for your companionship. Thanks to Shirley, Jacky and Ann from the Music Department for helping me out with administration issues.

Thanks to my participants for their interest in participating to my studies. Thank you to Edward Large and other peers for their advice on publications related to this work. Thank you to my examiners Prof Uwe Grimm, and Prof Eduardo Miranda whose work on brain-computer music interfacing has certainly set the current work in motion from as early as 2002.



# Contents

1	INTRODUCTION . . . . .	1
2	HUMAN RHYTHM PERCEPTION . . . . .	9
2.1	The temporal structure of stimuli . . . . .	10
2.1.1	Metricality . . . . .	11
2.1.2	Temporal nuances . . . . .	14
2.2	Beat perception . . . . .	15
2.2.1	Beat as mental pulsation . . . . .	15
2.2.2	Multiplicity of beat perception . . . . .	16
2.2.3	Dynamic perception of beat . . . . .	17
2.3	Meter perception . . . . .	19
2.4	Polyrhythm perception . . . . .	21
2.4.1	Polyrhythmic stimuli . . . . .	21
2.4.2	Human tapping behaviours in polyrhythm perception . . . . .	22
2.5	Summary . . . . .	26
3	COMPUTATIONAL MODELS OF BEAT AND METER PERCEPTION . . . . .	27
3.1	Internal Clock Models . . . . .	28
3.2	Probabilistic models . . . . .	30
3.3	Oscillator Models . . . . .	33
3.4	Filter Array Models . . . . .	35
3.5	Issues of Physiological Plausibility . . . . .	36
3.6	Types of input encoding: symbolic and/or acoustic . . . . .	38

3.7	Causal and non-causal models . . . . .	39
3.8	Summary . . . . .	40
4	NEURAL RESONANCE MODEL . . . . .	41
4.1	Physiological plausibility of the neural resonance model . . . . .	42
4.2	Historical background in modelling a neural oscillator . . . . .	46
4.3	Mathematical derivation of the neural resonance model . . . . .	48
4.3.1	Mathematical derivation of a neural oscillator . . . . .	50
4.3.2	Canonical derivation of a neural oscillator . . . . .	54
4.4	Properties of neural oscillator . . . . .	58
4.4.1	Spontaneous oscillation . . . . .	58
4.4.2	Entrainment . . . . .	60
4.4.3	Higher Order Resonance . . . . .	60
4.5	Simulation of rhythm perception with the neural resonance model . . . . .	62
5	METHODOLOGY . . . . .	69
5.1	The methodology of validation . . . . .	70
5.2	Simulations (Tests) . . . . .	71
5.2.1	Input encoding . . . . .	72
5.2.2	Frequency range of the bank of oscillators . . . . .	74
5.2.3	Number of oscillators . . . . .	77
5.2.4	Duration of simulation . . . . .	78
5.2.5	Connectivity and number of networks . . . . .	80
5.2.6	Numerical solvers and nominal values of parameters . . . . .	81
5.3	Experimental method . . . . .	82
5.3.1	Participants . . . . .	83
5.3.2	Polyrhythmic stimulus and presentation rates . . . . .	83
5.3.3	Experimental task . . . . .	84
5.3.4	Stimulus presentation . . . . .	85

5.3.5	Method of comparing modelled with empirical data	85
5.3.6	Justification of selected method of comparison . . .	87
6	VALIDATION OF THE MODEL AGAINST EXISTING EMPIRICAL DATA . . . . .	89
6.1	Preparing to compare human with model behaviour . . .	90
6.2	Simulations with isochronous polyrhythmic stimuli . . . .	93
6.2.1	Simulation with a 4:3 polyrhythmic stimulus repeating once per second . . . . .	93
6.2.2	Discussion on simulation . . . . .	96
6.2.3	Simulations using the 4:3 polyrhythm at different tempi . . . . .	98
6.2.4	Simulation with a 3:2 polyrhythmic stimulus repeating once per second . . . . .	107
6.2.5	Simulation with a 2:5 polyrhythmic stimulus repeating once per second . . . . .	110
6.2.6	Simulation with a 3:5 polyrhythmic stimulus repeating once per second . . . . .	112
6.2.7	Simulation with a 4:5 polyrhythmic stimulus repeating once per second . . . . .	113
6.3	Generalising the frequency response pattern . . . . .	114
6.4	Simulations with non-isochronous polyrhythmic stimuli .	115
6.4.1	Preparation of polyrhythmic stimulus comprising non-isochronous pulse trains . . . . .	117
6.4.2	Stimulating the model . . . . .	119
6.5	Summary . . . . .	123
7	MODEL ANALYSIS . . . . .	125
7.1	Parameter sensitivity analysis . . . . .	126

7.1.1	Initial conditions of each oscillator: amplitude and phase . . . . .	127
7.1.2	Driving variable parameters: external polyrhythmic stimulus . . . . .	131
7.1.3	State variable parameters . . . . .	134
7.2	Cross-model comparison with the Wilson-Cowan model .	148
7.2.1	Frequency response pattern of the Wilson-Cowan model to a 4:3 polyrhythmic stimulus with 1 Hz repetition frequency. . . . .	150
7.2.2	Frequency detuning in the neural resonance model	151
7.2.3	Superimposed frequency response patterns for both the neural resonance and the Wilson-Cowan model . . .	152
7.3	Summary . . . . .	153
8	MODEL PREDICTIONS ABOUT HUMAN BEHAVIOUR . . . . .	155
8.1	Predictions to extend the categories of human tapping behaviours associated with polyrhythm perception . . . .	156
8.2	Prediction of the timescale people start to tap to polyrhythms . . . . .	157
8.2.1	Measuring relaxation time . . . . .	158
8.2.2	Modifying the original hypothesis . . . . .	160
8.2.3	Method limitations in evaluating the original hypothesis . . . . .	161
8.3	Experimental studies in human perception of polyrhythms	161
8.3.1	Experimental design and participants . . . . .	162
8.3.2	Data management and visualisation . . . . .	162
8.3.3	Raw data . . . . .	163
8.3.4	Identifying the category of tapping behaviour: an example . . . . .	165

8.3.5	Creating a dataset for each combination, i.e., polyrhythm type and tempo . . . . .	169
8.3.6	Visualisation of data using boxplots . . . . .	170
8.3.7	Empirical results of tapping along with a 4:3 polyrhythm for different tempi . . . . .	173
8.4	Evaluation of the model's predictions . . . . .	178
8.4.1	Extension of categories of tapping behaviours . . . . .	178
8.4.2	Evaluating oscillator relaxation time . . . . .	179
8.5	Summary . . . . .	190
9	CONCLUSIONS, LIMITATIONS AND FURTHER WORK . . . . .	193
9.1	Conclusions . . . . .	193
9.1.1	Validation against existing empirical data . . . . .	195
9.1.2	Evaluation of model's predictions . . . . .	196
9.2	Limitations . . . . .	198
9.3	Further Work . . . . .	199
9.4	Final Reflection . . . . .	204
	BIBLIOGRAPHY . . . . .	205



# List of Figures

Figure 2.1	Binary-branched representation of a metrical structure.	17
Figure 2.2	Categories of periodic tapping behaviours along with a 4:3 polyrhythm.	23
Figure 4.1	The anatomy of a neuron	44
Figure 4.2	Schematic representation of the neural oscillator.	45
Figure 4.3	Schematic representation of the neural oscillator.	49
Figure 4.4	A limit cycle of a neural oscillator exhibiting self-sustained oscillation (spontaneous oscillation).	52
Figure 4.5	Trajectories of spontaneous and damped oscillations.	59
Figure 4.6	An example of a neural resonance model that consists of an array of oscillators with frequencies from 0.5Hz - 8.0 Hz.	61
Figure 4.7	Average frequency responses of the neural resonance model in the presence of a series of metronomic clicks.	64
Figure 5.1	Amplitude response of oscillators to a 4:3 polyrhythmic stimulus.	79
Figure 6.1	Time series representation of a 4:3 polyrhythmic stimulus.	94

Figure 6.2	Frequency response of the neural resonance model in the presence of a 4:3 polyrhythm.	96
Figure 6.3	Response of a series of linear oscillators in the presence of a 4:3 polyrhythmic stimulus that repeats once per second. The polyrhythm consists of two pulses 4 Hz and 3 Hz.	98
Figure 6.4	Frequency response of the model in the presence of a 4:3 polyrhythm repeating once every 2400 msec.	100
Figure 6.5	Frequency response of the model in the presence of a 4:3 polyrhythm repeating once every 2000 msec.	101
Figure 6.6	Frequency response of the model in the presence of a 4:3 polyrhythm repeating once every 1800 msec.	101
Figure 6.7	Frequency response of the model in the presence of a 4:3 polyrhythm repeating once every 1600 msec.	102
Figure 6.8	Frequency response of the model in the presence of a 4:3 polyrhythm repeating once every 1400 msec.	102
Figure 6.9	Frequency response of the model in the presence of a 4:3 polyrhythm repeating once every 1200 msec.	103
Figure 6.10	Frequency response of the model in the presence of a 4:3 polyrhythm repeating once every 800 msec.	104
Figure 6.11	Frequency response of the model in the presence of a 4:3 polyrhythm repeating once every 600 msec.	104
Figure 6.12	Frequency response of the model in the presence of a 4:3 polyrhythm repeating once every 400 msec.	105
Figure 6.13	Frequency response of the model in the presence of a 3:2 polyrhythm repeating once every 1000 msec.	110
Figure 6.14	Frequency response of the model in the presence of a 2:5 polyrhythm repeating once every 1000 msec.	111

- Figure 6.15 Frequency response of the model in the presence of a 3:5 polyrhythm repeating once every 1000 msec. 112
- Figure 6.16 Frequency response of the model in the presence of a 4:5 polyrhythm repeating once every 1000 msec. 114
- Figure 6.17 Time series representation of a 4:3 polyrhythm performed by a human at 60 bpm. 118
- Figure 6.18 Frequency response of the model in the presence of a 4:3 polyrhythm performed by a human at 60 bpm. 120
- Figure 6.19 Time series representation of a 4:3 polyrhythm performed by a human at 60 bpm with no metronomic guidance. 122
- Figure 6.20 Frequency response of the model in the presence of a 4:3 polyrhythm performed by a human at 60 bpm with no metronomic guidance. 123
- Figure 7.1 Amplitude response of the neural resonance model over a period of 12.5 seconds where the initial amplitude for all oscillators is 0.7. 129
- Figure 7.2 Amplitude response of the neural resonance model over a period of 12.5 seconds where the initial amplitude for all oscillators is close to 0. 130
- Figure 7.3 Superimposed frequency response patterns for different initial amplitude of the oscillators 0.5 (blue) and near 0 (green). 130
- Figure 7.4 Superimposed frequency response patterns of the neural resonance model for the three different ways of encoding a polyrhythmic stimulus. 132

- Figure 7.5 Superimposed frequency response patterns for different amplitudes of a 4:3 polyrhythmic pattern with 1 Hz repetition frequency. 133
- Figure 7.6 Frequency response patterns for ( $\alpha = 0$ ) (blue), and ( $\alpha = -0.1$ ) (green). 141
- Figure 7.7 Frequency response patterns corresponding to sensitivity analysis for bifurcation parameter ( $\alpha$ ) and nonlinearity parameter ( $\varepsilon$ ). 143
- Figure 7.8 Frequency response patterns for ( $\alpha > 0$ ) (blue) and ( $\alpha = 0$ ) (green). 144
- Figure 7.9 Frequency response patterns corresponding to sensitivity analysis for saturation parameter ( $\beta$ ) and nonlinearity parameter ( $\varepsilon$ ). 147
- Figure 7.10 Frequency response patterns of the Wilson-Cowan model in the presence of a 4:3 polyrhythmic stimulus with 1 Hz repetition frequency for a range of different amplitudes of the stimulus. 151
- Figure 7.11 Frequency response pattern of the neural resonance model in the presence of a 4:3 polyrhythmic stimulus with 1 Hz repetition frequency for a negative value of the frequency detuning parameter ( $\delta$ ). 152
- Figure 7.12 Superimposed frequency response patterns of Wilson-Cowan and neural resonance model. 153
- Figure 8.1 Amplitude responses of the bank of oscillators in the presence of a 4:3 polyrhythmic stimulus with 1 Hz repetition frequency. 160
- Figure 8.2 Excerpt of pooled raw empirical data 164

- Figure 8.3 Timing data for successive taps in the presence of a 4:3 polyrhythmic stimulus repeating once every 1000 msec. 166
- Figure 8.4 An example of a dismissed trial. 167
- Figure 8.5 Empirical tapping data 169
- Figure 8.6 Multiple boxplot summarizing the categories of tapping and the range of time it takes for people to start tapping along with the given 4:3 polyrhythm repeating once every 1000 msec. 171
- Figure 8.7 Descriptive statistics for the 4:3 polyrhythm repeating once per second. 172
- Figure 8.8 Multiple boxplots for the range of tapping modes and their corresponding period of time within which people start tapping along with a 4:3 polyrhythm repeating once every 2500 msec. 173
- Figure 8.9 Multiple boxplots for the range of tapping modes and their corresponding period of time within which people start tapping along with a 4:3 polyrhythm repeating once every 2000 msec. 174
- Figure 8.10 Multiple boxplots for the range of tapping modes and their corresponding period of time within which people start tapping along with a 4:3 polyrhythm repeating once every 1800 msec. 174
- Figure 8.11 Multiple boxplots for the range of tapping modes and their corresponding period of time within which people start tapping along with a 4:3 polyrhythm repeating once every 1600 msec. 175

- Figure 8.12 Multiple boxplots for the range of tapping modes and their corresponding period of time within which people start tapping along with a 4:3 polyrhythm repeating once every 1400 msec. 175
- Figure 8.13 Multiple boxplots for the range of tapping modes and their corresponding period of time within which people start tapping along with a 4:3 polyrhythm repeating once every 1200 msec. 176
- Figure 8.14 Multiple boxplots for the range of tapping modes and their corresponding period of time within which people start tapping along with a 4:3 polyrhythm repeating once every 800 msec. 176
- Figure 8.15 Multiple boxplots for the range of tapping modes and their corresponding period of time within which people start tapping along with a 4:3 polyrhythm repeating once every 600 msec. 177
- Figure 8.16 Multiple boxplots for the range of tapping modes and their corresponding period of time within which people start tapping along with a 4:3 polyrhythm repeating once every 400 msec. 177
- Figure 8.17 Temporal development of resonances among the bank of oscillators in the presence of a 4:3 polyrhythm repeating once per second. 185
- Figure 8.18 The period of time it takes for people to start tapping in a given mode along with a 4:3 polyrhythm repeating once every 1000 msec. 186

- Figure 8.19 The modes of tapping and their corresponding periods of time (boxplots) within which people start tapping along with a 3:2 polyrhythm repeating once every 1000 msec. 188
- Figure 8.20 Temporal development of resonances in the network in the presence of a 3:2 polyrhythm repeating once every 1000 msec. 188
- Figure 8.21 The modes of tapping and their corresponding periods of time (boxplots) within which people start tapping along with a 2:5 polyrhythm repeating once every 1000 msec. 189
- Figure 8.22 Temporal development of resonances in the network in the presence of a 5:2 polyrhythm repeating once every 1000 msec. 189
- Figure 9.1 Frequency response of a network of interconnected oscillators in the presence of a 4:3 polyrhythmic stimulus of 1 Hz repetition frequency. 202

## List of Tables

- Table 2.1 Metrical hierarchy representation 21
- Table 2.2 Tapping rates along with different polyrhythms 25

Table 5.1	Potential periodic human tapping behaviours along with a 4:3 polyrhythm expressed in frequencies for a series of repetition rates. 76
Table 6.1	Periodic tapping behaviours along with a 4:3 polyrhythmic stimulus expressed in terms of frequencies, when the repetition frequency of the polyrhythmic pattern is 1 Hz. 92
Table 6.2	A generic frequency response pattern of the neural resonance model in the presence of a two-pulse train polyrhythm. 114
Table 7.1	Range of numerical values for parameters ( $\alpha$ ), ( $\beta$ ) and ( $\varepsilon$ ). Values are calculated based on a $\pm 10\%$ uniform variation around nominal values. 139
Table 7.2	Combinations of different runs for parameters ( $\alpha$ ) and ( $\varepsilon$ ) considering both single and multiple sensitivity analysis. 142
Table 7.3	Combinations of runs regarding parameters ( $\beta$ ) and ( $\varepsilon$ ) considering both <i>single</i> and <i>multiple</i> parameter analysis. 146





## INTRODUCTION

---

People dancing, musicians playing music together and concertgoers tapping their feet while watching a band perform are all examples of behaviours which are mediated through the perception of musical rhythm. Various theories have been suggested to account for human rhythm perception, some informed by musicological analysis, others in which rhythm perception is approached as an engineering task, and yet more where the concern is to understand the physiological underpinnings that give rise to the perception of rhythm. The perception of rhythm is a fundamental catalyst for the creation and appreciation of the art of music. Research shows that listening to music triggers the generation of emotions and also, that engagement with rhythmic music coordinates brain activity which can both play a profound role as a therapeutic medium for various psychological and biological diseases associated with brain malfunction (Thaut et al., 2005). Understanding the physiological underpinnings of rhythm perception therefore can shed light to the above areas and promote human well-being in a non-invasive way. The theory of neural resonance has relatively strong physiological plausibility and rather than simply matching overt features of human rhythm perception, it shows how these features might plausibly emerge from low-level properties of neurons. As a result, the theory of neural resonance can be used as a medium to promote our understanding in how the brain perceives rhythm and inform the design of non-invasive ways of promoting healthy brain functionality.

In this thesis we focus principally on one theory of rhythm perception, namely the theory of neural resonance (Large and Snyder, 2009) according to which key aspects of musical experience [such as rhythm perception] can be explained directly in terms of brain dynamics (Large, 2008). The physiological plausibility of the neural resonance theory, in combination with its scope and competency (which are described in detail later) make it an excellent candidate theory of rhythm perception. The aim of this thesis is to evaluate the theory of neural resonance by examining the extent to which the theory can provide an account for human perception of a complex class of rhythms known as polyrhythms. Thus, the research question in general terms can be stated as follows.

To what extent can the theory of neural resonance provide an account for human perception of polyrhythms?

This question can be broken down into two halves.

Question 1: To what extent can the theory of neural resonance provide an account for existing empirical data on how humans perceive polyrhythms?

Question 2: Can we make predictions about human behaviour in polyrhythm perception based on the theory of neural resonance and subsequently test these predictions?

In order to address the above research questions, a four-fold set of research objectives will be outlined below. But first of all, to help illuminate these objectives, we will quickly sketch the bare bones of neural resonance theory. Briefly, neural resonance theory suggests that there are a number

of different populations of neurons in the brain where each population behaves like an oscillator each with its own natural frequency. In the presence of an external rhythmic stimulus some of these populations react to the stimulus and a pattern of multiple responses is established. The theory proposes that this pattern is responsible for giving rise to the perception of rhythm. The neural resonance theory may be instantiated as a computational model using a bank of oscillators, where each oscillator simulates a single neural population. The particular computational model used in this thesis (Large et al., 2010) (described and justified in detail in Chapter 4) will from now onwards be referred to as the 'neural resonance model'.

Various tests have been already carried out on the model (Large, 2010), to provide empirical support for the neural resonance theory with respect to key aspects of musical experience including aspects of not just rhythm, but also tonality perception. The objective of this thesis goes beyond existing tests of the theory by conducting a series of trials that examine the extent to which the neural resonance theory can account for aspects of human polyrhythm perception. Thus the thesis deals with a class of rhythms that have not been previously studied in the context of the theory and the model.

Polyrhythms refer to a particular subset of a class of rhythms where at least two different rhythms are present in parallel. In its simplest instantiation a polyrhythm can be created by superimposing at least two different streams of pulses where each one has a different periodicity. In existing experimental studies carried out by music psychologists, polyrhythms are considered an advantageous rhythmic stimulus in comparison to other simple rhythms (Handel, 1984) because they balance one of the more complex forms of rhythmical structure to be encountered in authentic musical pieces with ease of experimental control. Using

polyrhythms to test the neural resonance model inherits similar advantages: it provides an opportunity to use stimuli that encompass some of the complexity of authentic musical contexts, while allowing carefully graded comparisons to be arranged.

Additionally, the use of polyrhythms to test the neural resonance model is particularly advantageous because it bypasses limitations of the model (or any comparable model) in dealing with melodic or harmonic elements of a stimulus. The multiple rhythms of a polyrhythm afford an alternative way to represent the full range of conflicting rhythmical cues in real stimuli conveyed by several physical characteristics of the stimulus such as pitch.

The objective of the thesis can now be presented broken down into four sub-objectives:

- 1) Conduct tests to examine whether the neural resonance model can provide an account for existing empirical data on how humans perceive polyrhythms.
- 2) Conduct model analysis to limit the sources of uncertainty with regard to the physiological plausibility of the assumptions of the theory/model and to examine the dependency of the model on the numerical values of its parameters.
- 3) Use the neural resonance model to make predictions about human behaviour in polyrhythms previously unobserved.
- 4) Conduct novel experimental studies in human polyrhythm perception in order to evaluate the predictions made in 3.

Briefly, Chapter 2 contains background information on the phenomenon of human rhythm perception with an emphasis on fundamental aspects to be accounted for in modelling attempts. Chapter 3 reviews a range of computational models of rhythm perception that exist in the literature. Chapter 4 is devoted to introducing the neural resonance model in detail. Chapter 5 discusses the general methodological position taken in this thesis and it describes the methods adopted in Chapters 6 and 8. Chapters 6-8 contain original contributions. Chapter 6 validates the neural resonance model against existing empirical data, Chapter 7 focuses on analysing the neural resonance model and Chapter 8 evaluates specific predictions about human behaviour made by the model. Chapter 9 is devoted to conclusions, limitations and future work.

More specifically, Chapter 2 discusses the role of the structure of an external stimulus in eliciting rhythm perception, as well as the challenges it imposes in modelling attempts. Polyhythmic stimuli are introduced as the main type of stimulus used in this thesis. Also, it introduces beat and meter perception as the main aspects of human rhythm perception in order to highlight particular characteristics of the phenomenon that any theory or model of rhythm perception should address. Chapter 3 reviews a range of computational models of rhythm perception, with an emphasis on underlining some of the modeling dimensions, such as different computational methods and scope. Chapter 4 begins by discussing the physiological plausibility of the neural resonance model, it then introduces its mathematical formulation and level of abstraction, and reviews studies whereby the model is used to simulate human rhythm perception. Chapter 5 is concerned with the methodology and methods of this thesis. It begins by briefly discussing the notion of validation in order to indicate appropriate ways of interpreting the results obtained by the studies

of this thesis. It continues by describing the method of setting up the neural resonance model for running simulations to obtain modelled data. The method of conducting experimental studies in human polyrhythm perception is described following predictions about human behaviour made by the model. The chapter finishes by explaining and justifying the adopted method of comparison between modelled and empirical data in this thesis, which is employed in Chapters 6 and 8.

Chapter 6 focuses on answering the first half of the research question based on evidence provided by utilising sub-objective 1 in the above list. In particular, the chapter is concerned with the outcome of a series of tests in the model using polyrhythmic stimuli and how these outcomes compare with existing empirical data. The outcome of the above comparisons is the type of evidence used to address the first half of the research question, which is to examine the extent to which the theory of neural resonance can provide an account for existing empirical data on how humans perceive polyrhythms.

In particular, the type of human behaviour that the model aims to account for is the ways people tap along with a polyrhythm in a periodic manner, a behaviour similar to that of concertgoers tapping their foot along with a performance. The study is conducted on the basis of five different types of polyrhythms each one presented at ten different tempi, and empirical data are borrowed from the study of Handel and Oshinsky (1981) in human polyrhythm perception. Lastly, the outcome of the simulations suggests that the way the neural resonance model responds in the presence of a two-rhythm polyrhythm can be generalised, which may be used as a reference to inform and update the general consensus on how humans perceive polyrhythms.

Chapter 7 is concerned with tackling sources of uncertainty arising from the modeling assumptions of the neural resonance model and the nominal numerical values given in the parameters of the model (sub-objective 2). In particular the method of parameter sensitivity analysis is employed with regard to examining the sensitivity of the model's behaviour given a particular set of numerical values. Uncertainty with regard to biological assumptions is tackled by employing the method of cross-model comparison, whereby the behaviour of the neural resonance model is compared with the behaviour of another widely accepted biological model regarding neural dynamics, namely the Wilson-Cowan model (Wilson and Cowan, 1972). The neural resonance model is consistent with the way it responds regardless the systematic alteration in its numerical values and also, it produces similar responses to the Wilson-Cowan, thus the results presented in Chapter 6 and 8 can be considered robust. To our best knowledge this is the first study in which parameter sensitivity analysis is applied to the neural resonance model.

Chapter 8 introduces predictions made by the neural resonance model about human behaviour related to polyrhythm perception. In particular, the behaviour concerned here is the time it takes for people to start tapping along with a polyrhythm. These predictions drive the design of novel experimental studies in human polyrhythm perception. Briefly, 37 participants were recruited to take part in the study. The experimental variables monitored the time it takes for people to start tapping to polyrhythms, and the mode of tapping, i.e., tapping frequency. The empirical results were subsequently compared with the model's predictions. The model manages to successfully predict the time it takes for people to start tapping up to a certain degree. Additionally, the empirical data have shown new ways of how people tap (or perceive beat) along



with polyrhythmic stimuli. The time it takes for people to start tapping along with a range of polyrhythms is considered an experimental variable, which according to our best knowledge is monitored for the first time in the literature. Chapter 8 contains the type of evidence defined by sub-objectives 3 and 4 and thus, it is used to answer the second half of the research question, which examines whether the model can make predictions about aspects of behaviour related to human perception of polyrhythms.

## HUMAN RHYTHM PERCEPTION

---

A fundamental aspect of the phenomenon of human rhythm perception is the perceptual identification of periodicity in the presence of a rhythmic stimulus. There are many everyday behavioural manifestations of this phenomenon, for example when audiences clap their hands in time with a song. In fact, as we will see later, one of the principal methodologies for the experimental investigation of rhythm perception is to ask people to tap periodically along with rhythmic stimuli. Rhythmic periodicity can be characterised in terms of beat and meter, the two principal perceptual aspects of rhythm perception. Beat and meter are briefly introduced below, before a more comprehensive treatment in Sections 2.2 and 2.3.

An intuitive way of thinking about *beat perception* is to think of it as the feeling of an ongoing, steady, and unaccented pulsation in the presence of a rhythmic stimulus. *Meter perception* corresponds to the feeling of a regular pattern of strong pulses interposed by a series of weak pulses. Research in the area of beat and meter perception employs empirical methods to illuminate various facets of both processes. In this chapter we describe fundamental aspects of both beat and meter perception that any theory of rhythm perception should be able to account for.

In Section 2.1 we begin by discussing the role of the structure of stimuli in eliciting rhythm perception. That is, beat and meter perception arise in the presence of stimuli that exhibit certain objectively measurable structures. Section 2.1 introduces these structural characteristics of

rhythmic stimuli before we start analysing the accompanying perceptual phenomena. Sections 2.2 and 2.3 duly move on to consider fundamental aspects of beat and meter perception. Section 2.4 is concerned with polyrhythms, the particular class of rhythmic stimuli considered in this thesis. Section 2.5 summarises the chapter.

## 2.1 THE TEMPORAL STRUCTURE OF STIMULI

The way humans identify periodicity in the context of listening to stimuli depends on their temporal structure. More specifically, the temporal structure of stimuli determines whether a rhythm can be perceived or not, and constrains the various ways beat and meter may be perceived. In other words, the temporal structure of a stimulus can be thought of as a temporal matrix that defines the points in time where a stimulus event has to occur. For the purposes of this thesis, we refer to this structure as metrical because of its organised form and we provide a detailed description of that form in Subsection 2.1.1. In Subsection 2.1.2 we consider some possible nuances that may be present in a metrically organized structure in order to describe more precisely the actual way rhythmic information is conveyed to and perceived by humans. Before we continue we should note that metricality refers to objectively measurable structural characteristics and dissociate it with the act of perceiving meter, which is the result of listening to some stimulus with a metrical structure. In this section we deal with the first while meter perception is addressed in Section 2.3.

### 2.1.1 *Metricality*

The main characteristic of the organisation of a metrical structure is the co-existence of multiple levels of periodicities that comply with particular constraints (disclosed below) (Patel et al., 2005). The above set of different periodicities are related to each other in such way that over time the metrical structure behaves like a repeating pattern (London, 2004). A metric pattern can first be labeled on the total number of elements or time points it involves. The time points are those points in time that define the periodicities. The aforementioned total number of time points usually relates to the smallest period present in the structure, and it is the total number of points encountered in one repetition of the metric pattern. London refers to this label as the  $N$ -cycle of the metric pattern where  $N$  is the number of points in the cycle. Given that all other periods in the structure can be expressed as multiples of the smallest one, we can then express all periods as an  $(N/n)$ -cycle where  $(n)$  is the number of times we have to multiply the smallest period to get another one. For example a  $(N/2)$ -cycle corresponds to a period twice the length of the smallest period and it has  $(N/2)$  points. According to London (2004) cyclical representations of metrical structure follow five well-formedness constraints that arise due to specific psychological and cognitive limits in humans. The constraints listed below correspond to the underlying constraints a metrical structure complies with.

- WFC 1: The inter-onset intervals (IOIs) between the time points on to the  $N$ -cycle (and all sub-cycles) must be categorically equivalent. That is, they must be nominally isochronous and must be at least 100 msec approximately.

- WFC 2: Each cycle - the N-cycle and all subcycles - must be continuous, that is, they must form a closed loop.
- WFC 3: The N-cycle and all subcycles must begin and end at the same temporal location, that is they must all be in phase.
- WFC 4: The N-cycle and all subcycles must all span the same amount of time, that is, all cumulative periods must be equivalent. The maximum span for any cycle may not be greater than 5 seconds approximately.
- WFC 5: Each subcycle must connect nonadjacent time points on the next lowest cycle. For example, a level with a period defined by a quarter note has a period twice as long in comparison to a period defined by an eighth note in another level. That is every quarter note connects two nonadjacent eighth notes.

An example of a stimulus that complies with the above constraints is a drumbeat whereby a drummer plays 16th notes on the hi-hat and quarter notes on the bass drum at the same time. There are two different periods defined by the durations of the 16th and the quarter notes. The relationship between 16th notes and quarter notes remains invariant over time that is four 16th notes are repeated for every quarter note. Also, the onset of the quarter note coincides (at least approximately) with the onset of the first 16th note.

For the purposes of this thesis we will revisit WFC 5 in order to explore the types of ratios between periods from different levels that can be encountered in a metrical stimulus. It is useful to note that some authors have used 'metrical' and 'polyrhythmic' as mutually exclusive terms, but in line with other authors, and for the purposes of this thesis, we will consider polyrhythm to be simply a special case of metrical rhythm.

Amongst other things, this has the considerable benefit that we can refer to London's well-formedness constraints while discussing polyrhythms. Polyrhythms are described in detail in the next section. In order to distinguish polyrhythms from metrical rhythms that are not polyrhythmic, we need to revisit WFC 5 and discuss its implications regarding the possible ratios between two periods from different levels.

WFC 5 implies that two periods from two different levels relate to each other by standing in a ratio such as 2:1, 3:1, 4:1, etc. For example, a 4:1 ratio implies that a period of one level is 4 times longer than the period of another level. This matches the earlier example of a drummer playing 16th notes on the hi-hat on top of quarter notes on the bass drum. Another example would be three triplet eighth notes (quavers) on top of one quarter note, representing a 3:1 ratio.

Trivially, any ratio between two integers may be interpreted as defining a rational number. So for example the ratios 2:1, 3:1, 4:1 may be interpreted as defining the rational numbers  $2/1$ ,  $3/1$ ,  $4/1$  respectively. Clearly, these fractions happen to reduce to the integers 2, 3 and 4. However, there are other ratios commonly found between periods in rhythms that correspond to rational numbers that cannot be reduced to integers, for example 2:3 and 4:3 ratios. However, such ratios in no way hinder the perception of beat or meter and for that reason we consider them as valid ratios that can be found in a metrical structure. For example, the duration of a triplet quarter note is  $2/3$  of the duration of a standard quarter note, i.e., three triplet quarter notes have the same duration as two standard quarter notes. Ratios of this kind are very common in polyrhythmic stimuli, as discussed in detail in Section 2.3.

### 2.1.2 *Temporal nuances*

The notion of periodicity discussed above does not need to imply strict isochronicity. WFC 1 states that successive periods in one level should be categorically equivalent and nominally isochronous. That is, a series of events associated with one level of periodicity does not need to occur absolutely periodically. For example, the periods defined by the 16th notes played by a drummer will deviate from being perfectly isochronous, however they are considered to be categorically equivalent. Furthermore, in musical performances musicians employ purposeful temporal fluctuations in their playing in order to convey musical meaning. The above type of deviation from perfect periodicity in the underlying structure of the stimulus is also known as expressive variation - what musicians might call "pushing and pulling the time" (London, 2004, p.28). Another temporal nuance in the structure of a stimulus can be observed in complex musical rhythms whereby not all points in time - where an event onset is expected to occur - are actually employed by a stimulus event. For example, in syncopated stimuli events occur in some positions in the metrical structure while leaving nearby positions empty in a non-regular way (Fitch and Rosenfeld, 2007). According to (Large, 2008), both temporal deviation from perfect isochrony and syncopation constitute forms of complexity that are the most troublesome for modelling rhythm perception.

This completes our survey of objectively measurable features of stimuli that afford beat and metre perception. In the next section we turn to the associated topic of analysing the accompanying perceptual phenomena. The importance of the next section is to discuss fundamental aspects of both beat and meter perception that any theory that attempts to account for rhythm perception should account for.

## 2.2 BEAT PERCEPTION

In this section, three attributes of beat perception are considered. These are expressed in a statement form as follows:

- Beat takes the form of a perceived pulsation in the mind of the listener,
- Beat is perceived subjectively, i.e., it is not an explicit part of the stimulus and it can be felt in a number of different ways,
- Beat is dynamic in its nature, and constantly influenced by the stimulus.

### 2.2.1 *Beat as mental pulsation*

In the introduction of this chapter it was mentioned that beat perception corresponds to the identification of an ongoing, steady, and unaccented pulsation that is attributed to a rhythmic (or metrical) stimulus. Cooper and Meyer (1960) describe beat perception as 'a series of regularly recurring, precisely equivalent' psychological events that arise in response to a metrical stimulus. Cooper and Meyer's view in seeing beat as a perceptual construct rather than a property of the stimulus is commonly shared by music theorists, such as London (2004), and Lerdahl and Jackendoff (1983). The dissociation between perception of beat and the stimulus further indicates a human ability to feel pulsation independently of the presence of a stimulus, thus it further supports the view of beat as a perceptual construct, which is felt subjectively. In fact, empirical studies in tracking complex sequences of metrically structured rhythms have shown that humans can synchronise to complex rhythms in different subjective ways (Large et al., 2002). The behavioural manifestation of the above type of synchronization is tapping behaviours that correspond to different levels



of periodicities where the latter may or may not be marked by a stimulus event (e.g., syncopated stimuli)(Large et al., 2002).

### 2.2.2 *Multiplicity of beat perception*

The variability of human behaviour in perceiving a beat given a particular stimulus is a main characteristic of beat perception. A brief example of this kind of behavioural variation in beat perception and its link to the metrical structure of a particular stimulus is spelled out below.

Figure 2.1<sup>1</sup> below uses the example of a simple monophonic melody (labelled 'surface level') consisting of a series of note events to illustrate the multiplicity of beat perception. In this figure, the different levels of beat perceived by different people are explicitly indicated. In concrete terms, each different level in the hierarchy may be thought of as an indication of one possible way of tapping one's foot along with a song. This characterization in terms of 'foot tapping' has its limits: level n-2 is probably too fast to tap a foot to, given that it corresponds to 16th notes. The middle level (hierarchical level n) corresponds to quarter notes with respect to the provided notation at surface level and it most likely represents the most popular level for people to tap to. Where any metrical structure has a beat level that is the most popular for people to tap to, this level is referred to as the *tactus* (Lerdahl and Jackendoff, 1983). Drake et al. (2000) note that while individuals may gravitate to one particular beat level, however, they can generally change to either higher or lower levels if they wish (e.g., by doubling or halving the tapping rate) and still feel synchronised with the music.

---

<sup>1</sup> Figure 2.1 derived from Drake et al. (2000, p. 253)



Figure 2.1: Binary-branched representation of a metrical structure. The vertical lines delineate potential levels of beat perception in the presence of the rhythmical stimulus at the surface level

### 2.2.3 *Dynamic perception of beat*

The perception of beat helps individuals to create expectations about when events are likely to occur in the future. We could think of that process as the principal mechanism that allows people to play music together or to dance to a song. However, it should be stated that these expectations are dynamic in their nature and constantly influenced and challenged by the presence of a stimulus. That is, in order for those expectations to be retained and fulfilled, the stimulus should exhibit relative temporal stability over the course of its presentation, a stability that is nonetheless flexible enough to accommodate minor temporal fluctuations related to the event onsets of a stimulus. As was earlier discussed in Subsection 2.1.3 temporal fluctuations (particularly local nuances) occur naturally in the structure of a stimulus performed by humans, e.g., expressive variation. Rankin, Large, and Fink (2009) have found that the fluctuation patterns of expressive piano performance (by means of analyzing the time series of the performance) are

structured, and have demonstrated that people are still able to perceive a steady beat in such performance. That is, the perception of beat is flexible enough to retain its stability even in cases where the stimulus exhibits small deviation from perfect periodicity.

For a more extreme example, an empirical study by Large, Fink, and Kelso (2002) used a series of clicks with temporal fluctuations chosen purposefully to be larger than the fluctuations normally found in expressive variation. The clicks featured tempo perturbations of plus and minus 15 % around a central inter-onset interval of 400 msec. This study showed that humans can successfully re-synchronise tapping within 5 clicks after the perturbation and without having to stop tapping during the adaptation phase. In general, people are able to adapt to both tempo and phase perturbations of simple and complex rhythmic patterns (Large et al., 2002; Thaut et al., 1998; Repp, 2002). Additionally, in cases where individuals have control over the production of a stimulus, its temporal structure can be altered on demand and consequently, the perceived beat is also changed. For example, musicians can influence the ongoing perceived periodicity either by gradually accelerating (*accelerando*) or decelerating (*ritardando*) their playing as part of the stylistic execution of a musical piece, and still be able to perceive a beat. The aforementioned types of adaptability cast light on the dynamic nature of beat perception, that is, the perception of beat is anticipated based on the recently past events, and that changes in the presentation of a stimulus constantly influence and challenge the perception of the beat. In general, we can think of the beat as a dynamic act of identifying a pulse that runs through a metrical stimulus.

The discussion so far has highlighted several characteristics of beat perception such as its dynamic and subjective nature, and also, the multiplicity of beat perception, i.e., different ways of attributing a steady

beat to a given stimulus (see Figure 2.1). Some of the key features of beat perception can be summarised in the words of Meyer and Cooper (1960, p. 3):

‘Though generally established and supported by objective stimuli (sounds), the sense of pulse (beat) may exist subjectively. A sense of pulse, once established, tends to be continued in the mind and musculature of the listener, even though the sound has stopped. For instance, objective pulses may cease or may fail for a time to coincide with the previously established pulse series. When this occurs, the human need for the security of an actual stimulus or for simplicity of response generally makes such passages point toward the re-establishment of objective pulses or to a return to pulse coincidence’.

Having considered some of the key features of the perception beat, we will now consider the key features of meter perception.

## 2.3 METER PERCEPTION

In meter perception there is a sense of grouping of a series of successive beats. This is marked by the perception of certain periodically recurring beats as louder or more emphasised compared to the rest in the series. Those beats that are perceived as louder are generally perceived as the first in a pattern. Meter perception therefore can be described as a feeling of regularly positioned strong beats called metrical accents (Lerdahl and Jackendoff, 1983) interposed by a series of weak pulses. This strength is not necessarily associated with phenomenal accents, e.g., *sforzandi*, sudden changes in dynamics or timbre, or long notes (Lerdahl and Jackendoff,

1983), but rather the issue is one of perceived loudness. A common example of perceived metrical accents is the phenomenon of subjective accentuation whereby an isochronous series of notes of identical frequency and intensity are grouped in a pattern of two or four notes with the first note subjectively perceived as louder (Bolton, 1894). Bolton further notes that the recurring of accented sounds forms a secondary rhythm out of the primary rhythm. This secondary rhythm is sub-harmonically related to the main rhythmic frequency. That is, meter perception may be viewed as an instance of the ability of humans to spontaneously hear subharmonics of a rhythmic frequency presented to them, a characteristic that repeatedly arises in the way people perceive polyrhythms, as we will see in Section 2.4.

In general, given a metrical stimulus, metrical accents will arise in temporal locations where the beats of many different levels of periodicities come into phase (Large, 2000). In other words, a time point which is perceived as a beat on two different levels of pulsation is 'structurally stronger' than a point which is felt as a beat in only one level (Clayton et al., 2005) p.31. Thus, the perception of meter is based not only on the existence of several levels of periodicities, but also on the fact that different levels come into phase, which was previously described as one of the constraints in the way different levels of periodicity relate to each other (see Subsection 2.1.2 - WFC<sub>3</sub>). The more beats from different levels come into phase, the stronger structurally is that point perceived to be. That means that if a stimulus event occurs at a point in time of a structurally stronger beat then it may be perceived as louder than if it occurs at a structurally weaker beat. Figure 2.2 below is Lerdahl and Jackendoff's classic representation for illustrating structural strength and its association with phenomenal loudness. For example, events occurring at beat position 1 would be felt more strongly than events occurring at beat position 3,

and events occurring at beat position 2 would be felt weaker than events occurring at beat position 3 or 1.

Table 2.1: Lerdahl and Jackendoff's diagrammatic representation of metrical hierarchy based on several beats.

Beats	1	2	3	4	1	2	3	4
Level 1	•	•	•	•	•	•	•	•
Level 2	•		•		•		•	
Level 3	•				•			

## 2.4 POLYRHYTHM PERCEPTION

In this section we introduce polyrhythmic stimuli in detail and we describe the different behaviourally overt ways in which people perceive beat and meter in a polyrhythm.

### 2.4.1 *Polyrhythmic stimuli*

In Section 2.1.1 we discussed how, in the case of metrical music that is not polyrhythmic, the ratios between the periods of different levels of beat reduce to integers (e.g. 2:1, 3:1 and 4:1). By contrast, polyrhythmic stimuli in their simplest form consist of two or more different isochronous pulse trains which are in phase, but the ratios of whose periods do not reduce to integers (e.g., 2:3, 3:4, etc). In more complex examples, one or more isochronous pulse trains might be replaced with a complex pattern or figure in such a way that the actual or implied isochronous pulse trains have periods with ratios that do not reduce to integers.

To take a simple example, a polyrhythm could consist of a 3-pulse train dividing the duration of a shared repeating pattern into three equal IOIs, while a 2-pulse train divides the duration of the pattern into two equal IOIs.

In Subsection 2.1.1 we have noted that for the purposes of this thesis, and in line with London's (2004) Well Formedness Constraints for metrical structures, polyrhythmic structure can be thought of as a special case of metrical structure. As previously described, a metrical structure contains different periodicities that co-exist in parallel and satisfy particular constraints, such as that different periods stay invariant over time, and the onsets of the periods come into phase regularly. Polyrhythmic structure complies with the above constraints and therefore it is considered metrical for the purposes of this thesis.

#### 2.4.2 *Human tapping behaviours in polyrhythm perception*

One of the simplest forms of a polyrhythm consists of just two rhythms (pulse trains), e.g., rhythms  $m$  and  $n$ . Just as with simple meters, different people may perceive different beats or meters in polyrhythms, or the same people may have different perceptions of the same stimulus on different occasions. According to (Pressing et al., 1996) the most frequently occurring ways of perceiving beat or meter in polyrhythms, as behaviourally manifested by tapping along in a periodic way, are as follows:

- (a) Tapping along with the  $m$  pulse train.
- (b) Tapping along with the  $n$  pulse train.

Additionally, other empirically observed modes of tapping along with a polyrhythm include (Handel and Oshinsky, 1981):

- (c) Tapping along with the points of co-occurrence of the two pulse trains (i.e., once per pattern repetition or measure).

(d) Tapping along with every second element of one of the pulse trains, e.g., every other element of an m-pulse train or an n-pulse train.

Figure 2.2 below illustrates the cases a to d of the above list, relabeled as cases 1 to 5, i.e., the two cases in d) are considered independently. To make the illustration concrete, a 4:3 polyrhythmic stimulus is used as an example in Figure 2.2.

**4 over 3 polyrhythm**

**Observed categories of human tapping behaviour in dotted format (from top to bottom)**

1. Tapping along with the coincidence of the two rhythms
2. Tapping along with the 4-pulse rhythm
3. Tapping along with every other element of the 4-pulse rhythm
4. Tapping along with the 3-pulse rhythm
5. Tapping along with every other element of the 3-pulse rhythm

Figure 2.2: Categories of periodic tapping behaviours along with a 4:3 polyrhythm. The two measures of the polyrhythmic pattern depicted above are sufficient to illustrate the cyclicity of the above tapping options.

According to Handel and Oshinsky (1981), 80% of the behaviourally observed responses - for a number of different polyrhythms and tempi - correspond to tapping along with the elements of one of the pulse trains. The second major class (12% of the observed behaviours) of tapping periodically along with a polyrhythmic stimulus in a periodic way is to tap in synchrony with the co-occurrence of the two pulses, or once per



measure (coincidence of pulse trains). The third class (6%) corresponds to humans tapping periodically along with every other element of a 4-pulse train (most common), and every other element of a 3-pulse train (least common). An influential factor that shapes the human tapping preferences is the repetition frequency of the polyrhythm. Empirical results from the Handel and Oshinsky study demonstrate that the choices made by listeners in the interpretation of a polyrhythms, as revealed by regular tapping, depend in part on the absolute tempo.

The empirical results reported by Handel and Oshinsky also reveal limits on the tapping speed that humans can exhibit. For example, in the case of a 4:3 polyrhythm, when the polyrhythm repeats once every 600 msec, the fastest tapping observed corresponds to tapping along with the 3-pulse train i.e. tapping once every 200 msec. The slowest tapping rate for a 4:3 polyrhythm repeating once per second corresponds to tapping once every 1000 msec (i.e. tap along with the coincidence of the two pulse-trains). The table below illustrates minimum and maximum tapping speeds for all of the different types of polyrhythms and tempi explored by Handel and Oshinsky. Tapping speeds range from one tap every 120 msec to one tap every 1200 msec.

Table 2.2: Fastest and slowest tapping speeds along with different polyrhythms reported by Handel and Oshinsky (1982).

Polyrhythm Type	Fastest Tapping	Slowest Tapping
3/4	200 msec (tapping along with the 3-pulse train when the polyrhythm repeats once per 600 msec)	1000 msec (tapping along with the coincidence of the two pulse trains when the polyrhythm repeats once per 1000 msec)
2/3	200 msec (tapping along with the 2-pulse train when the polyrhythm repeats once per 400 msec)	1200 msec (tapping along with the coincidence of the two pulse trains when the polyrhythm repeats once per 1200 msec)
3/5	160 msec (tapping along with every other element of the 5-pulse train when the polyrhythm repeats once per 400 msec)	800 msec (tapping along with the coincidence of the two pulse trains when the polyrhythm repeats once per 800 msec)
2/5	120 msec (tapping along with the 5-pulse train when the polyrhythm repeats once per 600 msec)	1200 msec (tapping along with the 2-pulse train when the polyrhythm repeats once per 2400 msec)
4/5	150 msec (tapping along with 4-pulse train when the polyrhythm repeats once per 600 msec)	600 msec (tapping along with the coincidence of the two pulse trains when the polyrhythm repeats once per 600 msec)

## 2.5 SUMMARY

Both beat and meter are perceptual mechanisms that involve the identification of a periodicity in association with the presence of a metrical stimulus. Certain aspects of beat and meter perception, such as the ability of humans to accommodate small temporal fluctuations (both local and wider) in the structure of a stimulus without disrupting the general feeling of an underlying periodicity have been identified as some of the principal aspects of the phenomenon of rhythm perception that any theory of rhythm perception should be able to account for. Lastly, given that polyrhythms have been characterised as metrical, different ways people perceive beat and meter in polyrhythms have been described in terms of regular tapping, and empirical data regarding how fast people tap along with polyrhythms have been summarised.

## COMPUTATIONAL MODELS OF BEAT AND METER PERCEPTION

---

In the previous chapter, various phenomena were itemized and described that any model of rhythm perception should be able to account for. Firstly, candidate models ought to be able to account for the perception of beat and meter as outlined in the previous chapter. Secondly, they should be able to deal with the resilience of beat and meter perception in the presence of various kinds of temporal variation, such as time-varying tempi, also discussed in that chapter. Thirdly, they should be able to deal with the full variety of types of metrical structures (in the sense of 'metrical' as defined in the previous chapter) including, for example polyrhythms. The above considerations can be applied both to theoretical and computational models of rhythm perception. Depending on the modeler's perspective, these considerations can be interpreted either as a series of modelling tasks that have to be implemented and incorporated into a model, or as a series of criteria that models should comply with.

In this chapter, we will review computational (as opposed to purely theoretical) models of rhythm perception. The review is representative rather than exhaustive. Models are categorised and critiqued in several ways. First of all, each model is critiqued in terms of the degree to which it is able to account for the various phenomena of rhythm perception that must be accounted for, as itemized above. Secondly, consideration is made of the extent to which each model is, or is not, physiologically plausible.

We consider four principal categories of models. In the list below, one or more implementations is cited against each category of model. Each of the four categories is discussed in detail in Sections 3.1 - 3.4.

3.1 Internal clock models (Povel and Essens, 1985)

3.2 Probabilistic models (Laroche, 2001; Seppänen, 2001)

3.3 Oscillator models (Church and Broadbent, 1990; Large and Kolen, 1994; McAuley, 1995; Toiviainen, 1998)

3.4 Filter Array Models (Oppenheim and Schafer, 1975; Scheirer, 1998; McAngus Todd et al., 1999; Eck, 2007; Klapuri et al., 2006)

Having critiqued each of the four categories as listed above, we then discuss three general emergent issues as follows.

3.5 Issues of physiological plausibility

3.6 Issues of input encoding in models

3.7 Causal vs. non causal models

### 3.1 INTERNAL CLOCK MODELS

Internal clock models are associated with some of the first attempts in modelling human perception of rhythm. The core idea of these models is that humans possess one or more internal clocks of some kind and some kind of way of effectively counting and storing (or at any rate estimating) the number of ticks between two events of interest. Models

vary according to whether there is assumed to be a single such clock, or multiple such clocks, and whether the tick period is assumed to be fixed (e.g., at 1 millisecond), or whether it may be possible for the tick period to be influenced by the period of external stimuli of interest. In some models, two or more clocks are assumed to be related hierarchically (Povel and Essens, 1985). Povel and Essens posit that two characteristics need to be defined to determine the setting of a clock. These are its unit length (tick length) and its starting time (i.e., the time that counting starts). The starting time of the clock is always defined relative to the time of some stimulus event and by default the two times are assumed to be the same. As already noted, in some variant models the unit length may be assumed to take a value based on the interval periods present in a stimulus (within certain limits).

Povel and Essens postulate that a whole family of clocks with different tick rates is judged ("scored") against any incoming stimulus. The judgment or score is based on the extent to which the ticks of any particular clock coincide, or fail to coincide, with the incoming events. The clock with the highest score is the most strongly favoured ("induced") clock and is chosen to represent the clock that humans possess in perceiving rhythm. Its temporal unit corresponds to a fundamental pulsation of the stimulus such as that of the *tactus*. The clock then calculates a second pulsation, which is the subdivision of its unit.

Because the model is hierarchical, it satisfies our criterion of being able to produce multiple responses capable of reflecting the metrical structure of a stimulus. Despite the fact that this model is considered hierarchical, the version of the model described in the Povel and Essens (1985) paper only captures two metrical levels, for example the periodicity related to the *tactus* level, and one subdivision. However, Povel and Essens claim that

the model can be extended to identify more of the levels of the metrical structure of a stimulus. But there is another, deeper limitation of this model. For example, given a polyrhythmic stimulus the model may be able to identify a pulsation of one pulse-train but then information with regard to the pulsation underlying another pulse-train in the polyrhythm would be undiscovered, because the period defined by one of the pulse trains in a polyrhythm cannot be an integer multiple of the period defined by another pulse train in the same polyrhythm.

With regard to the ability of the internal clock model to capturing the temporal nuances of a stimulus described in Chapter 2, Povel and Essens (1985) note that future developments of the model can allow for a clock that continuously adapts its unit length, that is it can deal with stimuli that exhibit temporal fluctuations with respect to perfect isochrony. That is the internal clock model has the potential to comply with one more of the criteria set in the beginning of this chapter.

Lastly, given Povel and Essens assumption that people have access to an internal clock, they implicitly identify their model as a model that is related to physiology. Despite the appeal to physiological processes implicit in Povel and Essens's assumption, the computational formalism of the model utilises optimisation methods - where a "best" clock is chosen among a series of candidate clocks to represent the perception of rhythm in a stimulus. This multiplicity of clocks, and the choice amongst them, is not explicitly linked to physiology.

### 3.2 PROBABILISTIC MODELS

In this section we review computational models that model rhythm perception using probability theory. The key idea is that successive

stimulus events define periodicities that can be analysed in order to predict when in the future an event is expected to occur. An example of a probabilistic model that processes audio files of actual musical stimuli is the Laroche model (2001). The above model is known as a 'beat model', which means that it explicitly focuses on identifying just one level of periodicity, e.g., the tactus. However, as noted in Chapter 2 the notion of beat perception is associated with the perception of multiple periodicities, so the description of the above model as a 'beat model' may be slightly misleading, because it identifies only one level of periodicity.

The computational process of the model in finding the beat is as follows. Initially, the model assumes that there is a constant tempo in a stimulus. It then identifies the first beat in a signal by means of transient analysis, i.e., times at which the energy of the signal in some frequency band increases sharply. Finally, the model defines a series of points in time at which there is a maximum probability to observe a beat (transient). The probability density function (PDF) around the projected beat times can be represented by any symmetrical distribution with a mean equal to zero, e.g., normal distribution. However, the variance of the distribution must be adjusted as a function of the tempo so the PDF does not spill into adjacent beats and does not become too small either.

Laroche's model is able to process audio stimuli that closely resemble actual musical stimuli however under the assumption that the tempo is constant. However, we have already described that it is very likely that musical stimuli would exhibit temporal nuances like tempo variation over time. In these cases Laroche mentions that dealing with time-varying tempi is much more difficult and that the algorithm should be improved.

Seppänen (2001) created a model of rhythm perception that analyses actual music stimuli at more than one metrical level. Similarly to



the Laroche model, Seppanen's model utilizes probabilistic methods in its processing. The model analyses a signal first at the tatum level (defined below), then at the tactus level, and finally, at all subordinate levels in between the tactus and the tatum level. For that reason it is more hierarchical than Laroche's single level analysis. The tatum level corresponds to the smallest metrical unit, that is the metrical level that resembles the fastest pulsation.

The processing of the stimulus begins by a transformation of the audio signal into a symbolic representation (see Section 3.4 for more on symbolic representation). An onset detector process is applied to the signal by means of tracking changes in the root-mean-square amplitude envelope of the signal and emits event onsets at points of rapid level increase. Then a process of estimating the tatum level is calculated on the basis of deriving the greatest common divisor of a number of inter-onset intervals (IOI). The IOIs are calculated by considering pairs of both adjacent and non-adjacent onsets. The model estimates the tatum level in real time based on the current signal information (see also Section 3.5). After the segmentation of the signal at the tatum level a psychoacoustic accentuation analysis takes place at each tatum onset in order to estimate the tactus (beat) level. The psychoacoustic accentuation is estimated by analyzing a number of signal characteristics such as onset power, onset spectral shape, bass level, etc. The beat is then estimated by observing the identified accentuated event onsets for periodicities near 100 BPM (beats per minute).

Seppanen's model allows for *accelerandos* and *ritardandos* in the stimulus thus it can accommodate stimuli that deviate from perfect isochrony. The only criticism with regard to the criteria we set out in the beginning of this chapter is that Seppanen's model does not account for all possible levels in a metrical structure since it is confined in identifying

levels between the tactus and the tatum level. Also, despite being a competent model for analyzing music like stimuli it is not explicitly related to physiological mechanisms underlying human rhythm perception.

### 3.3 OSCILLATOR MODELS

Oscillator models simulate rhythm perception by utilizing the generic properties of an oscillator such as its natural frequency, and its ability to adapt its phase and frequency (in the case of adaptive oscillators) in the presence of a rhythmic stimulus. Without using adaptive oscillators, the same effect can be obtained by using a set of oscillators.

For example, a set of oscillators can capture any level of periodicity in a metrical stimulus, as long as one oscillator has the corresponding natural frequency. Alternatively, a single oscillator with an adaptive frequency can do the same job. Similarly, temporal fluctuations in phase and frequency (either intentional or unintentional) can be tracked either by an adaptive oscillator or a set of oscillators.

Based on the generic properties of phase and period adaptation of an oscillator, different models that account for either beat or meter perception have been developed and evaluated in a number of situations. As early as 1984, Dannenberg developed a single oscillator model to identify beat in rhythmic patterns, which were either speeding-up or slowing-down (Dannenberg, 1984). Large and Kolen (1994) developed a model using a type of oscillator that was capable of tracking the beat in a complex piece of real improvised melodic performance. McAuley (1995) also created a model using a similar type of oscillator to the one used by Large and Kolen, which was able to follow large tempo changes (e.g., *ritardando*), thus the term adaptive oscillator was coined. McAuley (1995) evaluated his

model by employing a series of rhythmical patterns of varying complexity originally used by Povel and Essens (1985) to test their model and compared the output of his model with empirical results from people tapping along with the same patterns to eventually gain supportive results for his model. Also, Toiviainen (1998) created a model of beat tracking again based on a similar oscillator to the one used by Large and Kolen (1994). Toiviainen's model was designed to perform tempo tracking of a musical performance in real time, and hence it was used to creatively - and interactively - play back a predefined accompaniment in synchrony with the performance. Furthermore, according to Toiviainen (1998, p. 64) the model could track the beat in rhythmically complex inputs, such as melodies with trills, grace notes and syncopation, however, Toiviainen did not provide an accurate description of explicit examples of beat tracking performance.

A reasonable extension of the single oscillator models is to consider a network of oscillators rather than a single unit, in order to address the various levels of periodicity in the metrical structure of a stimulus. One of the first references in using a multiple-oscillator model to account for experimental results related to perception of time in rhythmical stimuli was reported by Collyer, Broadbent, and Church (1994). In particular, Collyer and colleagues asked participants to tap in an unconstrained manner along with a stimulus, and reported tapping preferences that followed a bimodal distribution with the two modes exhibiting inter-tap intervals of 272 and 450 msec. The two different ways of tapping along with the same stimulus were taken as indicative of the multiplicity of beat perception, and were eventually simulated by using the Church and Broadbent (1990) model of multiple oscillators. Miller, Scarborough, and Jones (1992) considered a network of coupled oscillators and observed its resonances to rhythmical patterns. A similar approach of interconnected units has been suggested by

Large and Kolen (1994) and further refined by (Large and Jones, 1999; Large, 2000; Large and Palmer, 2002; Large et al., 2010) where a bank of gradient frequency oscillators with filter-like behaviour has been developed and has been found to account for a series of challenging intricacies of rhythm perception such as stability of beat perception in the face of syncopated stimuli (Large, 2010). Furthermore, the oscillators used in the Large, Almonte, and Velasco (2010) model have a strong similarity to biophysical models of neural activity related to both single neurons and populations of neurons, as we will see. As a result, the theory of neural resonance has been proposed to account for rhythm perception (Large and Snyder, 2009). The theory suggests that when a network of oscillators spanning a range of natural frequencies is stimulated with a musical rhythm, a multi-frequency pattern of oscillations is established, where the latter accounts for a range of characteristics of the phenomenon of rhythm perception. Because of the tight link between the mathematical abstractions of this model and the physiology of neurons explored in more detail in Section 3.5 and in the next chapter, taken together with the high degree to which neural resonance models are able to account for the various phenomena of rhythm perception noted in the previous chapter, we will see that neural resonance models are good *prima facie* candidates for a physiologically plausible computational theory of rhythm perception.

### 3.4 FILTER ARRAY MODELS

The idea of having an array of filters to account for the perception of the metrical structure of a stimulus has also been explored by Oppenheim and Schafer (1975). Oppenheim and Schafer used an array of linear band-pass filters each of different frequency to identify the multiple periodicities

associated with the metrical structure of the stimulus. More recent models that utilize similar band-pass (or comb) filters include Scheirer's (1998) model of gradient-frequency bank of comb filters, Todd's et al. (1999) model of passband filters, Eck's (2007) autocorrelation model, and Klapuri's et al. (2006) model of resonating comb filters. Linear filters such as those described above have been quite successful in accounting for identification of beat and meter, in various musical styles. However, there are still some challenges remaining such as the inability to retain a stable beat in the face of syncopated stimuli and the time delay in capturing information about metrical structure on the level of the measure (Large, 2000). However, some of those models such as that of McAngus Todd, O'Boyle, and Lee (1999) are of a particular interest due to the effort they put into creating a "holistic" model of rhythm perception. That is they try to incorporate the empirically observed influence of motor production in rhythm perception and also they are concerned with how physical stimuli are processed neurobiologically. As Eric Clarke comments on an earlier version of Todd's model (McAngus Todd and Lee, 1994): "the model regards the tactus as the result of combining the output of the band-pass filter bank with a sensory-motor filter that has fixed and relatively narrow tuning characteristics and represents the intrinsic pendular dynamics of the human body" (Clarke, 1999).

### 3.5 ISSUES OF PHYSIOLOGICAL PLAUSIBILITY

Thaut et al. (2005) notes that by studying the physiology and neurology of brain function in music we can obtain a great deal of knowledge about general brain function, in regard to the perception of complex auditory stimuli, rhythm processing, language processing, and feeling of

emotion. Furthermore, Michon (1967) argues that mathematical models in and of themselves are psychologically void of sense if they are not somehow associated to physical and mental functions, like memory and order discrimination, which are known to play a role in time perception (Creelman, 1962; Sherrick Jr et al., 1961). For example, a central assumption in Michon's model of timing in temporal tracking, which focuses on simulating the way people perform in sensorimotor synchronization tasks, is that people are able to retain timing information in their memory about the duration of successive stimulus events and about the errors they perform in trying to produce a tapping behaviour that is in perfect synchrony with the onsets of the stimulus events. These two pieces of temporal information are drawn from memory in order to direct future behaviours with the aim to optimise the degree of synchrony between behaviour and event onsets. From that perspective, Michon's model is physiologically related, at least in principle, to mental functions like memory.

In general, models whose mathematical abstractions are tightly linked to the physiology of neurons are known as "neural" models. As noted in the previous section, one such model is the model of neural resonance. In general, the mathematical abstraction of such "neural" models is derived from well-established biophysical models of single neurons, e.g., Hodgkin and Huxley (1952), and networks or populations of neurons such as Wilson and Cowan (1972). Eck (2002) applied a relaxation oscillator model of neural spiking dynamics to the task of modeling beat induction in rhythmical patterns. This so-called relaxation oscillator has been used to model a range of biological oscillatory processes, including neural spiking (Hodgkin and Huxley (1952)) and heartbeat pacemaking (Van der Pol and Van der Mark, 1928). In fact, Eck's oscillator is a type of Fitzhugh-Nagumo

relaxation oscillator (Fitzhugh, 1961; Nagumo et al., 1962), which is a two-variable model of neural action potential, i.e., a simplification of the much more complicated model - but strongly biophysical - of Hodgkin and Huxley (1952). Eck's model has been tested using the rhythmical patterns suggested by Povel and Essens, and made good predictions in 34 out of the 35 cases. Also, the same tests were repeated after injecting noise (up to 20%) to the rhythmical patterns, thus to simulate natural temporal fluctuations in real stimuli. Despite the degradation of performance Eck reports an 82.9% successful rate in finding the beat.

### 3.6 TYPES OF INPUT ENCODING: SYMBOLIC AND/OR ACOUSTIC

A generic distinction among different models of rhythm perception, such as those discussed above, relates to the type of input that is fed into the model. In particular, the type of the input could either be an acoustical signal such as an audio file or a symbolic representation of a stimulus such as a MIDI file representing a musical score. Natural stimuli exist in the form of acoustic signals, so an argument has been made (Scheirer, 2000) contesting the validity of symbolic models in modelling human rhythm perception. Despite this, Seppänen (2001) notes that insofar models that process acoustic signals apply a signal-to-symbolic transform when processing the stimulus, there is not much sense in differentiating symbolic systems from acoustic signal processing systems. Additionally, there exist models, which are flexible in terms of the nature of the stimulus they consume. For example, Dixon (2001) proposes a symbolic MIDI tracker that can also consume acoustic signals. Similarly, the Large, Almonte, and Velasco (2010) neural resonance model provides great flexibility in encoding

the stimulus, where MIDI, acoustic and functional representations can all be employed to encode a rhythmical stimulus.

### 3.7 CAUSAL AND NON-CAUSAL MODELS

Another distinction between models of rhythm perception is whether a model is causal or not. The causality or non-causality of a model is based on whether the model's output requires considering the stimulus as a whole in advance of producing the output. In cases where the output of the model is determined on the basis of looking at the stimulus as a whole then the model is classified as non-causal, and causal otherwise. Causality is particularly important in cases where the aim is to produce a model that corresponds to the natural way people perceive rhythm, which is in essence a causal way, i.e., people start to perceive rhythm almost immediately after listening to a song. However, non-causality can be justified based on a hypothesis that humans would alter earlier percepts in retrospect, based on a later input, but only within the perceptual present of approximately 4 seconds (Parncutt, 1994). That is people have the ability to retain in their memory 4 seconds of a stimulus and perceive a rhythm retrospectively by processing those 4 seconds in their mind.

There are several models of rhythm perception that are causal. For example, in Scheirer's (1998) model the analysis of tempo and beat is performed causally. Toiviainen's (1998) interactive midi accompanist model tracks the tempo of a musical performance in real time. Large and Jones (1999) introduce their model as a dynamic model that builds on the notions of expectancy and entrainment to describe real-time tracking of time-varying events. Cemgil, Kappen, Desain, and Honing (2000)



formulate a tempo tracking system in a probabilistic framework that facilitates a reasonable first estimate of the presented stimulus.

### 3.8 SUMMARY

The discussion in Chapter 3 has provided an overview of a number of computational models of rhythm perception, such as clock-based, probabilistic, and oscillatory models. We have discussed the differences in their purposes, e.g., modelling either beat or meter perception, and pointed to limitations and advantages at least with regard to a number of criteria that are considered fundamental and challenging in modelling rhythm perception, such as tolerance to temporal variation in the structure of the stimuli, and capability of processing different forms of metrical structure.

As noted in Section 3.3 and Section 3.5, the neural resonance model shows a strong link between the mathematical abstractions and the physiology of neurons, and has good potential to account for the various phenomena of rhythm perception. It is not the only model for which claims of this kind can be made, but it is a good *prima facie* candidate for a physiologically plausible computational theory of rhythm perception. Consequently, the next chapter is completely devoted to the neural resonance theory and its model. The main focus is on describing the mathematical derivation of the model and its level of abstraction. Also, emphasis is given to understanding the physiological plausibility of the model in some detail. In order to do this, we will have to review some basic topics in brain physiology, and hypotheses about how this may relate to rhythm perception.

## NEURAL RESONANCE MODEL

---

In the previous chapter, the neural resonance model (Large et al., 2010) was identified as a physiologically plausible candidate model of rhythm perception, since it allows rhythm perception to be explained on the basis of brain dynamics (as explored in depth in this chapter). The neural resonance theory proposes that there are populations of neurons in the brain, and that each population behaves like an oscillatory system with its own macroscopic properties. These properties include synchronization of firing among the constituent neurons, and a collective natural frequency. Furthermore, when these populations of neurons are exposed to a regular stimulus, some of the populations start to resonate, i.e., start to oscillate with greater amplitude than in the absence of a stimulus. In physical terms, the amplitude of the oscillation corresponds to the amplitude of the electrical activity associated with the firing of a population of neurons. The neural resonance theory asserts that the collective resonance patterns formed in the presence of a metrical stimulus are responsible for giving rise to the perception of beat and meter. For example, the perception of a beat in a song emerges as a result of a population of neurons firing at a rate that creates the perceived beat.

In the first section of this chapter, the physiological and functional characteristics of the populations of neurons simulated by the model are described. We then introduce the neural resonance model in mathematical detail, while being careful to retain the link to physiological processes and

structures wherever possible. In the second part of the chapter we focus on existing studies where earlier versions of the neural resonance model were tested using various rhythmical stimuli, and in some cases validated against empirical data. Reviewing such studies is useful for framing the area of contribution of the present thesis, and also provides examples of the validation of computational models against empirical data. This is one of the validation methods employed in this thesis, as explained in detail in Chapter 5.

#### 4.1 PHYSIOLOGICAL PLAUSIBILITY OF THE NEURAL RESONANCE MODEL

The main unit in a population of neurons is a single neuron, and for that reason we will consider single neurons first. In particular, we focus on describing the physiology of neurons with particular attention to their action potential, which controls the production of discrete electrochemical signals. Neurons, together with glia cells, are the main cellular units that constitute the human brain. The number of neurons in an adult brain can be over 100 billion while the glia cells number about ten times more than that, with different concentration ratios at different parts of the brain. The neurons form networks of interactions by creating links called synapses. Synapses form the pathways through which neurotransmitters are passed from one neuron to the other. The formation of synapses is triggered by the propagation of an electrical pulse (action potential) within the neuron. The main parts of a neuron are dendrites (loosely speaking, inputs), soma (main body), and an axon (loosely, output) (Figure 4.1)<sup>1</sup>.

---

<sup>1</sup> Figure 4.1 derived from Arbib (2003), p. 1045

Action potentials propagate down the axon towards their terminals where they cause the release of varied neurotransmitters. The pulse-like propagation of action potentials allows cells whose elements are in continuous operation to function for many purposes as discrete processes. Thus the production of an action potential is also called firing of the neuron. When the neuron is at rest, the potential inside the cell is negative ( $-70\text{mV}$ ) relative to the outside of the cell. This potential is created by ATP-fuelled pumps in the wall of the neuron that consume one molecule of ATP to pump three sodium cations out for every two potassium cations pumped in. However, external stimuli, such as the reception of excitatory neurotransmitters from one or more excitatory presynaptic neurons (i.e., "input neurons") may cause ion channels to open in the axon membrane of the cell, allowing an influx of positive sodium ions ( $\text{Na}^+$ ), which causes depolarization, thus allowing the membrane potential to rise to values more positive than  $-70\text{mV}$ . At around  $-55\text{mV}$  there is a threshold that once it is reached it triggers the initiation of the action potential, i.e., a further rapid rise and then a fall in the potential. The maximum magnitude of the potential can reach  $100\text{mV}$ . By contrast, interaction with presynaptic inhibitory neurons (which are less well-understood) may for example decrease the permeability of the axon membrane to positive sodium ions leaking back in to the neuron, or may increase permeability to an influx of negative chloride ions ( $\text{Cl}^-$ ) and to an efflux of positive potassium ions ( $\text{K}^+$ ), any of which have a tendency to hyperpolarise the potential of the neuron. The fact that the neuron has many dendrites as inputs means that both excitatory and inhibitory activity can take place at the same time.

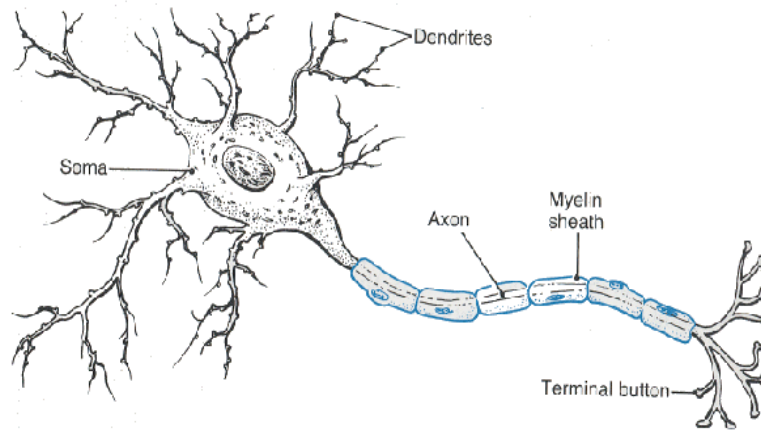


Figure 4.1: The anatomy of a neuron

Of particular interest for the purposes of this thesis is the functional characteristic whereby individual neurons exhibit regular periodic firing (Kalat, 2012). One such type of neuron is the thalamic reticular neuron: these have a pacemaker role in activating the sleep spindle rhythm - a type of firing activity that masks external stimuli during sleep in order to prevent sleep disruption (Wang and Rinzell, 1993). Additionally, not only single neurons but also ensembles of them of a particular local circuitry in various brain structures, such as the cerebellum, hippocampus, olfactory cortex, and neocortex have been found to produce periodic activity (Rakic, 1975; Shepherd, 1976). Hoppensteadt and Izhikevich (1996a, p. 118) explain that one of the basic mechanisms for periodic activity among a population of neurons is that local populations of excitatory and inhibitory neurons have extensive and strong synaptic connections so that action potentials generated by the former excite the latter, which, in turn, reciprocally inhibit the former. A neural population with functional characteristics such as periodic firing is called a neural oscillator. Physiological evidence for the existence of such populations (neural oscillators) is provided by the work of (Mountcastle, 1957) and (Hubel et al., 1963). Figure 4.2 illustrates

schematically a neural oscillator whereby one neuron from each type is depicted, i.e., excitatory and inhibitory. In practice, several thousands neurons of each type make up a typical neural oscillator<sup>2</sup>. These ensembles of neurons resemble the type of populations of neurons that the neural resonance model is concerned with. In particular the neural resonance model simulates the dynamics of such populations of neurons in the presence of a metrical stimulus.

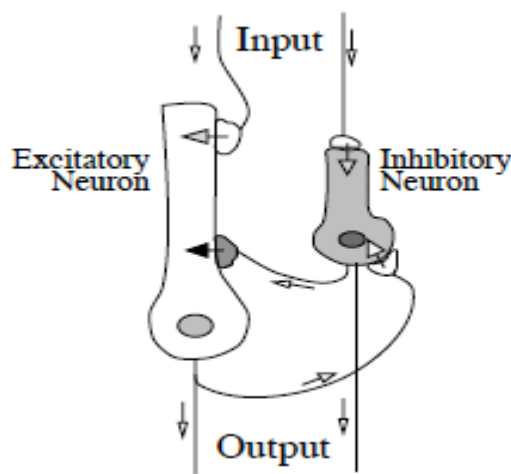


Figure 4.2: Schematic representation of the neural oscillator. It consists of excitatory (white) and inhibitory (shaded) populations of neurons. For simplicity one neuron from each population is pictured.

---

<sup>2</sup> Figure 4.2 derived from Hoppensteadt and Izhikevich (1997), p.118

## 4.2 HISTORICAL BACKGROUND IN MODELLING A NEURAL OSCILLATOR

The development of mathematical models to simulate activity of neurons is a long-standing process, with models focusing both on single neurons and populations of neurons. Also, there are different levels of mathematical abstraction in the development of these models, some being very detailed with respect to the physiology of neurons and others more abstract, where the interest is to capture properties (functional characteristics) of the neural activity. The neural resonance model belongs to the second category and it focuses on capturing the properties of a neural oscillator in the presence of a rhythmical stimulus, without its parameters being explicitly related to the physiology of neurons. However, its development is based on models that are quite detailed with respect to the physiology of neurons. In the next two paragraphs we will describe those models that have been influential in the development of the neural resonance model, which further allows us to illustrate models of both single neuron and populations of neurons using different mathematical abstractions.

One of the most popular biological models, Hodgkin and Huxley's (1952) model closely describes the physiology of the action potential of a single neuron. This model mimics the firing of single neurons as described in the previous section. The model has been derived from empirical observations of the electrodynamics of a living squid's axon. The electrodynamics of the neuron are modelled by using a system of non-linear differential equations that describe the behaviour of continuous currents on the membrane potential (Koch and Grothe, 2003). This model uses the differential equations to describe the initiation and propagation of action potentials in the axon. One drawback of this biological model is that

it requires detailed knowledge of many parameters, which additionally introduces a difficulty in defining all the degrees of freedom among the interaction of the parameters. Furthermore, being a single-cell model, it does not take into consideration the interaction of neurons within neural populations.

A seminal attempt to describe neural oscillations due to interactions of *populations* of excitatory and inhibitory neurons was the Wilson and Cowan (1973) model. It is based on the Fitzhugh (1961) and Nagumo, Arimoto, and Yoshizawa (1962) model of single-cell activity, which is a simplification, i.e. a more abstracted version, of the Hodgkin and Huxley (1952) model. The overall activity of the neural oscillator is modelled by attributing a differential equation to describe the activity of each given local population of either inhibitory or excitatory neurons (the equations have been omitted here as a qualitative description is sufficient for our purposes). The free parameters of the Wilson-Cowan model are highly biological. In the Wilson-Cowan's paper, 'A Mathematical Theory of the Functional Dynamics of Cortical and Thalamic Nervous Tissue' (1973) a long list of symbols in page 57 describes parameters such as post-synaptic membrane potential, surface density of excitatory neurons, and synaptic operating delay, all directly related to physiology and functionality of neurons.

The neural resonance model (Large et al., 2010) derives from the (Wilson and Cowan, 1972) model. In particular, the neural resonance model employs formal simplification methods that allow focusing on the *macroscopic* properties of neural oscillators as a system, and thus, it bypasses the need of an exhaustive list of physiological parameters and their interactions such as those in the Wilson-Cowan model. In fact, the neural resonance model is a 'canonical' model, a notion generally used to



refer to the simplest (in analytical terms) of a class of equivalent dynamical systems. In our case, the dynamical system is the neural oscillator itself. Other examples of canonical models that focus on neural oscillators are the Aronson et al. (1990) model and the Hoppensteadt and Izhikevich (1996a) model. The neural resonance model borrows the mathematical abstraction of a neural oscillator as developed in the Hoppensteadt and Izhikevich (1996a) paper, and extends it by assuming a network of heterogeneous frequency oscillators, i.e., oscillators with different natural frequencies, instead of homogeneous.

The formal mathematical introduction of the neural resonance model, its properties and its parameters are the subject of the next section. Regardless of the level of mathematical abstraction chosen to describe the phenomenon of neural oscillation, the mathematical properties of oscillators introduced in this section are universal, and thus expected to be observed in any mathematical approach related to oscillators.

#### 4.3 MATHEMATICAL DERIVATION OF THE NEURAL RESONANCE MODEL

This section begins with the mathematical representation of a single neural oscillator. The model comprises a collection of such oscillators, each one with its own natural frequency. In the mathematical expression of a neural oscillator that follows, Dale's principle (1935) is asserted throughout, whereby: the excitatory neurons may have only excitatory synaptic connections with other neurons and inhibitory neurons may have only inhibitory synaptic connections (Figure 4.3).

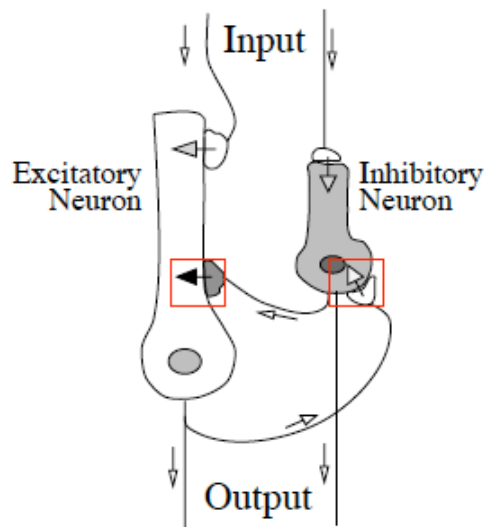


Figure 4.3: Schematic representation of the neural oscillator. It consists of excitatory (white) and inhibitory (shaded) populations of neurons. For simplicity one neuron from each population is pictured. Dale's principle is assumed, i.e., the white arrow in the red box denotes excitatory synaptic connection, while the black arrow denotes inhibitory synaptic connection. This simple coupling structure allows both excitatory and inhibitory modes to impose negative feedback on each other, thus moderating both extremes, and fostering steady oscillation of the pair.

#### 4.3.1 *Mathematical derivation of a neural oscillator*

The activity of a single neural oscillator (labelled by  $i = 1, 2, \dots, n$ ) can be thought of as a dynamical system and it can be modelled by a system of two differential equations, one describing the activity of an excitatory neuron and the other the activity of an inhibitory neuron (Equation 1). Each differential equation is a function of  $(x)$ ,  $(y)$  and  $(\lambda)$ .  $(x)$  is the activity of the excitatory neuron, and  $(y)$  is the activity of the inhibitory neuron. The index  $(i)$  is used to distinguish different oscillators, but it can be disregarded momentarily as we only deal with a single neural oscillator.  $\lambda$  is a bifurcation parameter, which for certain numerical values corresponds to a particular qualitative change (dynamical mode) in the properties of the neural oscillator.

$$\begin{aligned}\dot{x}_i &= f_i(x_i, y_i, \lambda) \\ \dot{y}_i &= g_i(x_i, y_i, \lambda)\end{aligned}\tag{1}$$

The externally measurable activity of excitatory and inhibitory neurons corresponds, at a macroscopic neurophysiological level, to the number of action potentials per unit time (Hoppensteadt and Izhikevich, 1996a, p. 119). Thus, the typical measure of neural oscillation in a given neighborhood is the varying potential monitored by a local electrode which detects the net difference between excitatory and inhibitory postsynaptic potentials in the neighborhood of the electrode (Wilson and Cowan, 1972). The parameter  $(\lambda)$ , the bifurcation parameter, defines the dynamical mode of the neural oscillator. For the purposes of this thesis, the focus is on the dynamical mode of the neural oscillator near an Andronov-Hopf bifurcation point (Andronov et al., 1971). Hoppensteadt and Izhikevich note that while an Andronov-Hopf bifurcation is only one of many possible bifurcations of

relevance for the dynamics of neural oscillation in general, it corresponds to a transition from equilibrium (resting) to periodic activity (firing). Thus, this mode is of most biological relevance for our purposes (Hoppensteadt and Izhikevich, 1996a), and is most relevant to the neural resonance model at the focus of this thesis (as considered in more detail in section 4.3.2).

Before we expand Equation 1 to deal with multiple neural oscillators, we briefly cover the idea of gaining information about the solution of a system of one neural oscillator by considering a geometric representation of the system's solution rather than an analytical solution. The reason for doing this is to introduce some useful notions that will facilitate the discussion as we go. Geometric methods in dynamical systems are quite common, especially when the differential equations governing the system are nonlinear and therefore arriving to a solution analytically is challenging. For example, Winfree (2001) shows how geometric methods of dynamics can be applied to the explication of biological oscillations, especially circadian and heart rhythms. In general, a solution of Equation 1 demands a mathematical expression of how the dependent variables ( $x$ ) and ( $y$ ) vary over time ( $t$ ). In order to understand a geometric representation of the solution we can define an abstract space with coordinates  $x$  and  $y$ .

The two-dimensional space (Figure 4.4) represents the phase space of the neural oscillator as a dynamical system. In general, the dimensions could be as many as the number of variables needed to characterise the state of the system, and thus, systems in general can be characterized as  $n$ -dimensional system or an  $n$ th-order system on that basis. Within the above space a trajectory can be drawn to correspond to the solution of the system over time in a particular mode and for given initial conditions. Interestingly, in many cases geometric reasoning allows us to draw the trajectories without actually solving the system (Strogatz, 1994). As an example, Figure 4.4

shows a trajectory that corresponds to a neural oscillator firing periodically (self-sustained oscillation), i.e., spontaneous oscillation. This trajectory is known as stable limit cycle. The limit cycle is a closed and isolated trajectory, which means that neighboring trajectories spiral either toward or away from the limit cycle (Strogatz, 1994, p. 196).

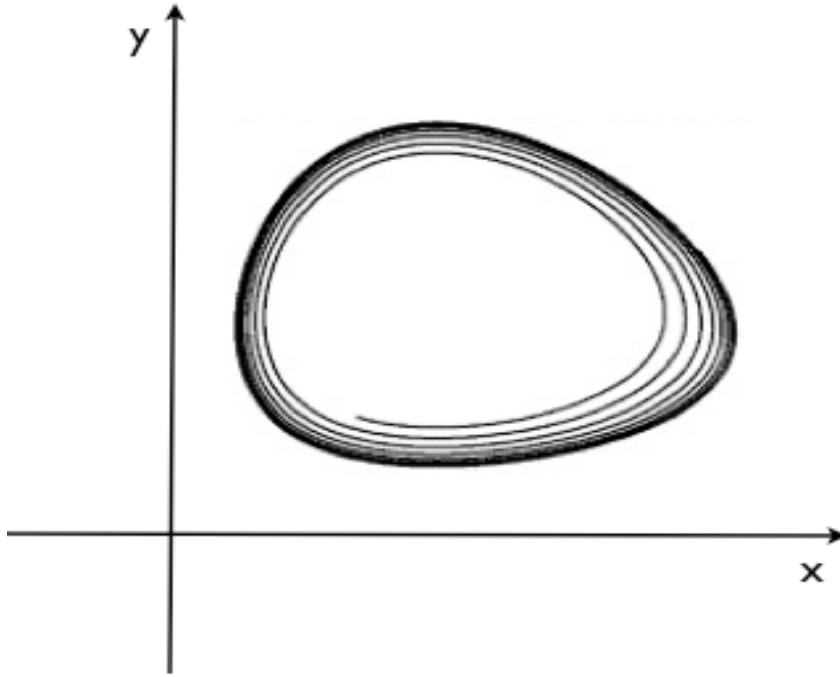


Figure 4.4: A limit cycle of a neural oscillator exhibiting self-sustained oscillation (spontaneous oscillation).

While a single neural oscillator can behave as outlined above, the neural resonance model is concerned with a network of gradient frequency oscillators. In order to model such an oscillator that is able to interconnect with other oscillators in a network, it is necessary to model its interactions with all the other oscillators in the network. In order to achieve this, we need to extend Equation 1 as follows. Equation 2 is a system of two differential equations describing the activity of a single neural oscillator that is coupled with other oscillators in a network. The first set of terms on the right hand side of the equations is the same as in Equation 1. The

second set of terms refers to the connection of a single oscillator with the rest of the oscillators in a network. Parameter ( $\varepsilon$ ) refers to the strength of the connection between oscillators.

$$\begin{aligned}\dot{x}_i &= f_i(x_i, y_i, \lambda) + \varepsilon p_i(x_1, y_1, \dots, x_n, y_n, \varepsilon) \\ \dot{y}_i &= g_i(x_i, y_i, \lambda) + \varepsilon q_i(x_1, y_1, \dots, x_n, y_n, \varepsilon)\end{aligned}\tag{2}$$

Here the functions  $p$  and  $q$  represent synaptic connections from the whole network of oscillators onto the  $i$ -th neural oscillator. The parameter ( $\varepsilon \geq 0$ ) represents the strength of synaptic connection between different neural oscillators. In general, parameter ( $\varepsilon$ ) is assumed to be small based on some neurophysiological justifications. More about the hypothesis of weakness of synaptic connections can be found in (Hoppensteadt and Izhikevich, 1995).

Equation 2 is of the right form to describe a neural oscillator weakly connected to all of the other neural oscillators in a network, but to make this equation useful for our purposes, we need to extend it again. The next step in extending the dynamical system is to consider its dynamics in the presence of an additional external input to the  $i$ -th oscillator. The external input resembles the presence of an external rhythmic stimulus of a pulse-like nature. Equation 3 is a system of two differential equations capturing the activity of a single neural oscillator in the presence of a rhythmic stimulus. The oscillator is coupled with other oscillators in the network. The parameters are the same as in Equation 2. The only new parameter is the external input  $\rho_{(x_i)}(t)$ , which is incorporated into the coupling term. Large, Almonte, and Velasco (2010, Equation 5) incorporate the external input in the following way:

$$\begin{aligned}\dot{x}_i &= f_i(x_i, y_i, \lambda) + \varepsilon p_i(x_1, y_1, \dots, x_n, y_n, \rho_{x_i}(t), \varepsilon) \\ \dot{y}_i &= g_i(x_i, y_i, \lambda) + \varepsilon q_i(x_1, y_1, \dots, x_n, y_n, \rho_{y_i}(t), \varepsilon)\end{aligned}\tag{3}$$

One can think of Equations 2 and 3 as being a generalization of the physiologically plausible Wilson and Cowan (1972) model of a neural oscillator in an oscillatory neural network (Hoppensteadt and Izhikevich, 1996a).

#### 4.3.2 *Canonical derivation of a neural oscillator*

In order to facilitate the analysis of a neural oscillator, we can focus on its state near the Andronov-Hopf bifurcation point. A focus on this particular bifurcation point is physiologically justifiable, because it resembles the states of resting and firing, as outlined in the previous section. This simplified focus also allows us to simplify the relevant equation by reducing it to a much simpler form (a canonical representation) through the application of normal form theory (Large et al., 2010). The exact derivation is omitted here, but a detailed procedure for obtaining normal forms of a system of equations like the one described by Equation 2 can be found in (Hoppensteadt and Izhikevich, 1996a; Wiggins, 2003; Kahn and Zarmi, 1998). While realistic mathematical models of physical and biological processes can be complicated (e.g., Hodgkin and Huxley's model of action potential), simplified forms can ease analysis while still capturing essential aspects of the phenomena being modelled.

The application of normal form methods in the system of differential equations described in Equation 2 simplifies it into a single equation for each  $(i)$  - rather than two equations for each  $(i)$  (Equation 4). However, variable  $(z)$  below is a complex valued state, which means that it has both

real and imaginary parts. Equation 4 below is a canonical representation of Equation 2.

$$\dot{z}_i = b_i z_i + d_i z_i |z_i|^2 + \sum_{j \neq i}^n c_{ij} z_j + \text{h.o.t.} \quad (4)$$

The meaning of the variables is momentarily withheld until we derive the fully expanded equation that incorporates an external stimulus and captures interaction between oscillators of different natural frequency. For example, the term h.o.t (higher order terms) captures the interactions between oscillators of different frequencies and it needs to be disclosed. As previously noted, Equation 3 considers the dynamical system of a neural oscillator in the presence of an external rhythmical stimulus. According to Large, Almonte, and Velasco (2010), the process of simplifying Equation 3 - by using the normal form method in a similar way as in Equation 2 - is problematical, and known methods for deriving a canonical model cannot be applied. In order to overcome this difficulty, Large et al. (2010) suggest that the external input  $s(t)$  to oscillator ( $z_i$ ) can be included in the coupling term as follows. Consider the coupling term - the second term on the right-hand side - of Equation 4, i.e.,  $\sum_{j \neq i}^n c_{ij} z_j$ . This term is related to the total amount of synaptic connection between a single neural oscillator and the rest of the oscillators in a network. The complex-valued coefficient ( $c_{ij}$ ) denotes the values of synaptic connections between neural oscillators, which depend upon ( $p_i$ ) and ( $q_i$ ) from Equation 2. In particular, the absolute value ( $|c_{ij}|$ ) encodes the strength of connection from the  $j$ -th to the  $i$ -th oscillators, whereas the argument of ( $c_{ij}$ ) encodes phase information (Hoppensteadt and Izhikevich, 1996a). The term  $x_i = \sum_{j \neq i}^n c_{ij} z_j + s(t)$  is the Large, Almonte, and Velasco (2010) suggestion for incorporating the external input into the coupling term, and as a result Equation 4 becomes



Equation 5. Note that the above  $(x_i)$  is not the same  $(x_i)$  as the one in Equations 1, 2, 3. The variables will be disclosed shortly.

$$\dot{z}_i = b_i z_i + d_i z_i |z_i|^2 + \sum_{j \neq i}^n c_{ij} z_j + s(t) + \text{h.o.t} \quad (5)$$

Equation 5 can be simplified into Equation 6 as follows:

$$\dot{z}_i = b_i z_i + d_i z_i |z_i|^2 + x_i(t) + \text{h.o.t} \quad (6)$$

Equation 6 is a canonical representation of a neural oscillator with an external stimulus incorporated into the coupling term. The higher order terms in Equation 6 may or may not have an effect on the dynamics of the system. We will treat these two cases in turn. In cases where the interacting neural oscillators have very close natural frequencies the effect of the higher order terms will be negligible (Large et al., 2010). This is the case considered by Hoppensteadt and Izhikevich (1996a) whereby a network of neural oscillators of homogeneous frequencies is assumed.<sup>3</sup> The neural resonance model we are considering in this thesis is a gradient frequency neural network, which comprises of a series of neural oscillators of different natural frequencies. In this case, the effect of the higher order terms in Equation 6 will not be negligible. Large, Almonte, and Velasco (2010) suggest that the development of a heterogeneous frequency oscillator network is of a critical relevance to understanding aspects of auditory processing, such as rhythm perception. In order to capture interactions between oscillators of different frequencies and also, responses of an oscillator to an input that does not share closely related frequency to

---

<sup>3</sup> According to Hoppensteadt and Izhikevich, oscillators of equal natural frequencies display functionally significant connections while oscillators of different natural frequencies display functionally insignificant connections (see Figure 4, Hoppensteadt and Izhikevich, 1996, p.120).

its natural frequency, the higher order terms need to be expanded. Equation 7 below is the fully expanded canonical representation of a neural oscillator ( $z_i$ ) that is part of a gradient frequency neural network that captures all the necessary interactions for modelling rhythm perception.

$$\dot{z}_i = b_i z_i + d_i z_i |z_i|^2 + \frac{k_i \varepsilon z_i |z_i|^4}{1 - \varepsilon |z_i|^2} + \frac{x_i(t)}{1 - \sqrt{\varepsilon} x_i(t)} \cdot \frac{1}{1 - \sqrt{\varepsilon} x_i(t)} \quad (7)$$

We are now in a position where the variables of the above equation can be disclosed. Parameters ( $b_i$ ), ( $d_i$ ), ( $k_i$ ), are all complex variables and they are defined as follows:

$$b = \alpha + i\omega$$

$$d = \beta_1 + i\delta_1$$

$$k = \beta_2 + i\delta_2$$

Equation 7 is two-dimensional, because  $z$  is a complex variable, having real ( $\text{Re}(z)$ ) and imaginary ( $\text{Im}(z)$ ) parts. It has both real ( $\alpha, \beta_1, \beta_2$ ) and imaginary ( $i\omega, i\delta_1, i\delta_2$ ) parameters, however ( $\omega$ ), ( $\delta_1$ ) and ( $\delta_2$ ) are all real. The parameters are ( $\alpha$ ), the bifurcation parameter, ( $\beta_1$ ) and ( $\beta_2$ ), the non-linear saturation parameters, ( $\omega$ ), the eigenfrequency of the oscillator, and ( $\delta_1$ ) and ( $\delta_2$ ), the frequency detuning parameters. The bifurcation parameter ( $\alpha$ ), determines the nature of the interaction between excitatory and inhibitory inputs (i.e., either spontaneous oscillation or damped oscillation - Figure 4.5). The saturation parameters ( $\beta_1$ ), and ( $\beta_2$ ) prevent amplitude from increasing indefinitely when  $\alpha > 0$  (Figure 4.5). The eigenfrequency ( $\omega$ ) is the natural frequency of the oscillator. The frequency detuning parameters ( $\delta_1$ ) and ( $\delta_2$ ) describe how the instantaneous frequency<sup>4</sup> of the oscillator also depends on its amplitude.

<sup>4</sup> Instantaneous frequency is  $d\phi/dt$  i.e., angular velocity

#### 4.4 PROPERTIES OF NEURAL OSCILLATOR

The remaining two sections of this chapter introduce and focus on those properties of a neural oscillator that account for aspects of rhythm perception. These are a) spontaneous oscillation, b) entrainment, and c) higher order resonance, which have been previously suggested (Large, 2008) to account for diverse phenomena, including: non-linear resonance in the cochlea; phase locked responses of auditory neurons; and entrainment of rhythmic responses in distributed cortical areas. Briefly, spontaneous oscillation refers to the activity of a neural oscillator that is *not* being stimulated by an external stimulus, i.e., self-powered oscillation. *Entrainment* describes the activity whereby the oscillator may adjust either its frequency or phase in the presence of some external rhythmical stimulus. *Higher order resonance* captures two distinct phenomena: firstly, interaction among multiple neural oscillators, each one with different natural frequency, and secondly, the response of each oscillator in the presence of a rhythmical stimulus with a pulse frequency that is *not* close to the natural frequency of that oscillator. The ways the above behaviours are mapped to the free parameters of the neural resonance model will be described below.

##### 4.4.1 *Spontaneous oscillation*

Spontaneous oscillation accounts for a dynamic state of the neural oscillator whereby periodic activity can be achieved independently of the presence of a rhythmic stimulus. Near an Andronov-Hopf bifurcation point, spontaneous oscillation is one of the potential dynamical states of a neural oscillator. Spontaneous oscillation occurs when its bifurcation parameter

is positive ( $\alpha > 0$ ). Figure 4.5 below illustrates how the amplitude of the oscillator stabilises in the spontaneous oscillation state<sup>5</sup>. In cases where ( $\alpha < 0$ ), the oscillator displays its second dynamical state, a damped oscillation, which means that the oscillation decays after some time (Figure 4.5). Bifurcation therefore can be thought of as the critical point whereby the system changes from one dynamical state to another. The bifurcation parameter responsible for controlling the transition between qualitatively different states of the oscillator is parameter ( $\alpha$ ) and the critical point occurs when ( $\alpha = 0$ ).

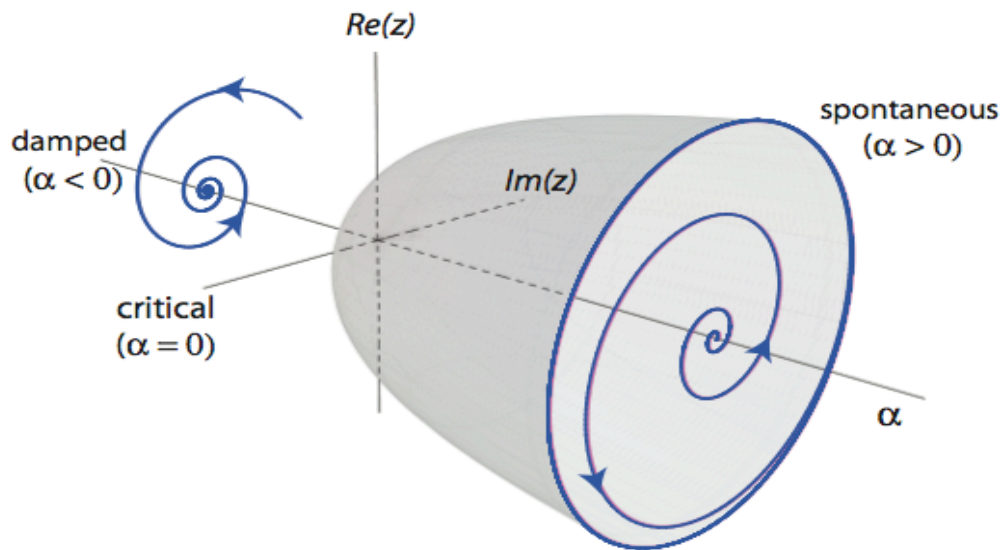


Figure 4.5: Trajectories of spontaneous and damped oscillations. For ( $\alpha > 0$ ), the amplitude of the oscillator increases up to a stable saturation point, while for ( $\alpha < 0$ ) the amplitude decays to zero.

We can see that, in the spontaneous oscillation mode, the oscillation reaches a particular amplitude, then stabilises, and repeats itself over time. This particular trajectory is an example of a limit cycle in dynamical systems. This dynamical state is of particular interest because it can be used to

<sup>5</sup> Figure 4.5 derived from (Large, 2008, p. 21)

simulate the observed periodicity of real neural oscillators (Section 4.1). In terms of rhythm perception, this property accounts for, and addresses, situations where the beat is felt even in the absence of stimulus events, e.g., pauses in a stream of note events.

#### 4.4.2 *Entrainment*

*Spontaneous oscillation*, by definition, is achieved independently of any external stimulus. In the case of *entrainment*, an established oscillation is coupled to an external stimulus of some frequency ( $f$ ). That is the oscillator develops a stable phase relation with the stimulus, i.e. phase locking. A typical example of phase locking is when both the oscillator and the external stimulus are constantly in phase, i.e., perfect synchrony. In phase locking can be observed only when both the oscillator and the external stimulus share the same frequency. If for example the frequency of an oscillator was twice as fast as the frequency of the stimulus then the phase locking would alternate between in-phase locking and anti-phase locking. Entrainment is central to neural resonance model in that the theory asserts that rhythm perception can be explained by the entrainment of physically real neural oscillators in the presence of a rhythmic stimulus.

#### 4.4.3 *Higher Order Resonance*

When the neural resonance model is considered *without* the higher order terms (h.o.t) it is only possible to model entrainment between a stimulus and an individual oscillator whose natural frequencies stand close to each other or exhibit simple integer ratios such as 2:1, 3:1, 4:1 (i.e. the frequency of the stimulus is more or less an integer multiple of the frequency of

the oscillator). The response of an oscillator that is not related to the stimulus in any of the above ways is inhibited. By contrast, in the fully expanded canonical model (Equation 7), where the higher order terms are fully expanded, any oscillator whose frequency stands in any low-integer ratio (or close approximation thereof) to the stimulus pulse frequency, for example not only the previously noted, but also 1:2 and 3:2 can resonate. For example, Figure 4.6<sup>6</sup> illustrates the responses of a network of neural oscillators in the presence of a rhythmic stimulus (a sinusoid of 2Hz). The responses of the network occur in a way that denotes sub-harmonic, harmonic and more complex ratios of the stimulus frequency. Apart from illustrating higher order resonance, Figure 4.6 addresses additional properties of a neural oscillator such as frequency detuning (e.g., high amplitude levels in the 1:1 ratio lead to a small mismatch between stimulus frequency and oscillator frequency), sensitivity to weak signals and high frequency selectivity, which have been previously observed as distinctive aspects of human rhythm perception (Large, 2008).

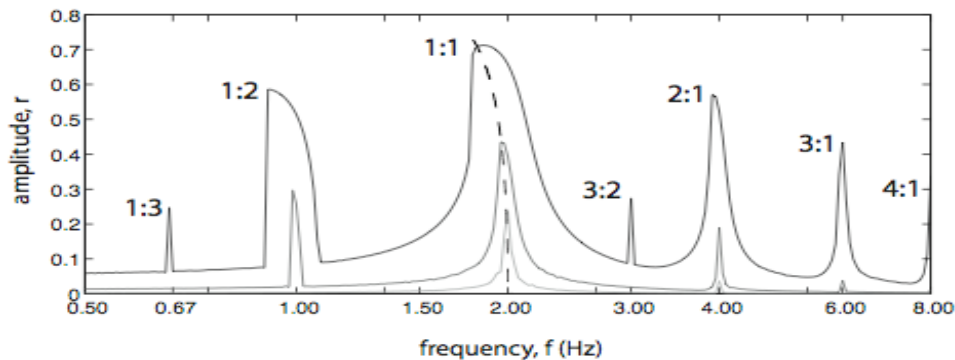


Figure 4.6: An example of a neural resonance model that consists of an array of oscillators with frequencies from 0.5 Hz - 8.0 Hz. The amplitude responses result from stimulating the model with a 2Hz sinusoid of various amplitudes.

<sup>6</sup> Figure 4.6 derived from (Large, 2008)

Aspects of the illustrated responses of the neural resonance model to external stimuli (Figure 4.6) play key roles in modelling the way humans perceive rhythm. For example, the multiple responses of the model to a 'monochromatic' stimulus suggests how the neural resonance theory might cast light on the ability of listeners to attend to multiple levels of the metrical structure of a stimulus under a wide variety of task conditions (e.g. playing an instrument). Figure 4.6 reveals that the strongest response is found at the stimulus frequency ( $f = 2\text{Hz}$ ), but responses are also observed at harmonics (e.g.,  $2f$ ,  $4f$ ) and sub-harmonics (e.g.,  $1f/2$ ) as well as rational ratios (e.g.,  $3f/2$ ) of the stimulus frequency. In other words, the responses of the oscillators in the presence of a stimulus may be able to account for human perception of metrical structure in simple rhythmic stimuli.

A strength of the neural resonance model is the demonstration of how higher-order interactions between oscillator and stimulus can give rise to stable oscillations and multi-frequency patterns of oscillation. These multi-frequency patterns of responses are very promising for testing the neural resonance model with more complex rhythmical stimuli such as polyrhythms, which is the main theme on this thesis.

#### 4.5 SIMULATION OF RHYTHM PERCEPTION WITH THE NEURAL RESONANCE MODEL

The theory of neural resonance has been proposed quite recently (Large, 2008; Large and Snyder, 2009) but computational implementations that utilise the idea of resonance to model human rhythm perception go back a number of years (Large and Kolen, 1994, 1998; Large and Jones, 1999; Large, 2000; Large and Palmer, 2002; Large, 2008). With this in mind, we review

cases where predecessor models of the one used in this thesis (Large et al., 2010) have been utilised to account for human rhythm perception using various stimuli such as simple metronomic clicks and stimuli that exhibit temporal variation and syncopation.

Figure 4.7 (below) illustrates an example of a resonant pattern produced by the neural resonance model in the presence of a simple rhythmical stimulus. In this case, the stimulus was a series of metronomic clicks with a tempo of 60 beats per minute (bpm) presented for a 100 second period of stimulation. In general, the resulting resonant pattern is dynamic, i.e., it evolves during the period while the metronomic clicks stimulate the bank of oscillators. However, the graph below corresponds to the average activity of each individual oscillator (blue dot) over the *second half* of the stimulation period (final 50 seconds). Averaging over the final part of the stimulation period ensures that appropriate time has been given for all oscillators to respond, and that the amplitude responses are measured only after the oscillators have reached a steady amplitude.

We can see that the neural resonance model exhibits peak responses, which are related in three different ways to the fundamental rhythmic frequency of the given stimulus: harmonically; sub-harmonically; and in a more complex fashion. In the case of Figure 4.7 a series of clicks played at 60 bpm imply a frequency of 1 Hz (i.e., one click per second). Consequently the fundamental pulse frequency of the stimulus is 1 Hz. We can see that there are peaks at 0.5 Hz, 1 Hz, 1.5 Hz, 2Hz, etc.

In terms of modelling human rhythm perception, some of the responses in Figure 4.7 can be interpreted in the following way. The highest peak in Figure 4.7, around the 1 Hz oscillator, corresponds in effect to listeners perceiving quarter notes (using the convention that the tempo of the clicks refers to quarter notes, i.e., ♩ = 60). Similarly, the peak around



the 2 Hz oscillator corresponds to listeners perceiving/playing eighth notes; the peak around the 1.5 Hz oscillator corresponds to listeners perceiving/playing triplet crochets; the peak around the 0.5 Hz oscillator corresponds to listeners perceiving/ playing half notes, and so on.

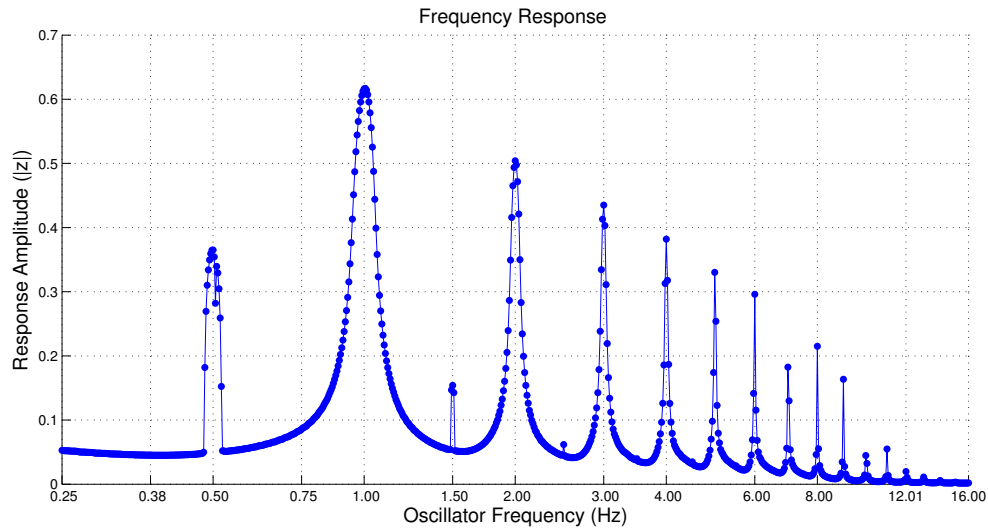


Figure 4.7: Average frequency responses of the neural resonance model after being stimulated by a series of clicks played at 60 bpm. The frequency of clicks is 1 Hz, i.e., one click per second. The highest peak corresponds to oscillators around the 1Hz oscillator, where the latter exhibits the maximum amplitude. Peaks related to oscillators with harmonically, subharmonically and more complexly related frequencies to the 1Hz oscillator can be observed. Some of these responses can be said to resemble the human perception of different metric levels associated with a given stimulus e.g., quarter notes (1 Hz), half-notes (0.5 Hz), 16th notes (4 Hz).

The above interpretation of Figure 4.7, using the neural resonance model illustrates how some of the principal aspects of rhythm perception, such as those described in Chapter 2, can be taken into account by the theory of neural resonance. For example, the model responds dynamically in the presence of a stimulus, and it identifies multiple levels of periodicity similar to those associated with human perception of beat and meter. The

metronomic clicks are nonetheless perfectly isochronous: the ability of the model to cope with temporal fluctuations will be examined in the following examples.

In Figure 4.7 the activated oscillators and their corresponding amplitude responses comprise the two main features of the neural resonance model's behaviour as a result of being stimulated by some simple periodic stimulus. Henceforth, for purposes of concise and clear presentation, the presentation of such data (Figure 4.7) is referred to as the frequency response pattern of the model.

Large and Kolen (1994) tested a model consisting of six oscillatory units of different natural frequencies covering a range of two octaves with a freely improvised melody performance collected in a study of musical improvisation conducted by Large, Palmer, and Pollack (1995). The performance displayed temporal variations at both local levels (expressive variation) and also rather large changes in tempo (*ritardando*). Two out of the six oscillators acquired stable modes when stimulated by the above performance as a stimulus, while the rest of the oscillators never stabilised. Based on the metrical structure of the performance as a stimulus the two stabilised oscillators were corresponding to the periodicity levels resembling the quarter note and the half note.

In another study (Large, 2000) investigated the response of a bank of oscillators similar to the one above in the presence of a series of musical excerpts of ragtime music as stimuli. Ragtime is a representative example of a syncopated type of stimulus, thus in effect the study was testing the ability of the model to identify a pulse in a syncopated stimulus. Syncopation amounts to a violation of the temporal expectations embodied in the perceived metrical structure and it can be operationalized as the *lack of events* on strong beats (or sometimes the occurrence of unstressed

events), accompanied by the placement of *stressed events on weak beats* (Large, 2000). A computer was used to generate musical events resembling a piano performance in a metronomically precise manner, i.e., expressive timing and tempo variation were restricted. This type of musical excerpt has been previously used by (Snyder and Krumhansl, 2000) as a means of identifying cues to pulse finding by humans tapping along with the ragtime examples. The effect of syncopation as a cue in pulse finding was evaluated by considering two versions of the performance. The first version was the full performance and the second one resembled the part of the performance that would have been played by the right hand, i.e., the syncopation in ragtime is usually performed by the right hand, while the left holds a steady bass rhythm. Large stimulated the model with both the full version and the syncopated version and he compared the responses of the model with human tapping behaviour reported by Snyder and Krumhansl (2000). In particular, he compared two aspects of human behavior with the behavior of the model. The first point of comparison was the mode of tapping, i.e., tapping frequency, and the second was the time it took for people to start tapping along with the ragtime excerpts. The human modes of tapping were compared with the most active oscillators presented by the model, and the time to start tapping was compared with the time it took for the corresponding oscillator to start resonating. The outcome of the two comparisons suggested that the simulation squares well with human performance (Large, 2000). Additionally, the outcome of the comparisons showed similarity between human behaviour and model response in terms of the degree of disruption caused by the level of syncopation. More specifically, when people were confronted only with the syncopated part of the performance more turbulence was observed in tapping behaviour - for example, switching from tapping the beat to tapping the off-beat - in

comparison to the full version. This behavioural finding was juxtaposed by Large (2000) with the model's deterioration in stabilizing to one beat in the presence of the syncopated part of the ragtime stimulus.

Another revealing example of pulse finding using the neural resonance model can be seen in the case of the highly syncopated son clave rhythmic pattern (Large, 2010). In particular, in the son clave rhythm note events occur at only half the points in time where the basic beat (tactus) is expected, and also, notes occur often on relatively weak beats, however, the neural resonance model identifies the underlying beat.

In this chapter, we have outlined the physiological plausibility of the neural resonance model by describing the physiological and functional characteristics of a particular type of population of neurons (neural oscillator) that the model aims to simulate. In particular, the model simulates the dynamics of these populations of neurons in the presence of a rhythmic stimulus. According to the theory of neural resonance, the characteristics of emerging multiple resonances give rise to rhythm perception. We have also described how these dynamics result from particular functional characteristics of oscillators such as entrainment, spontaneous oscillation, and higher order resonance. An extended mathematical description together with the level of mathematical abstraction of the neural resonance model has been provided. The development of the neural resonance model and its relation to various models has been briefly considered.

In Chapter 5, we will describe the process of setting up the neural resonance model for simulations to be run. Setting up the neural resonance model constitutes part of the general method for conducting validation tests with the model, preparatory to the comparison of the output of the model with empirical human data. Chapter 5 also introduces the method

of obtaining empirical human data and also the method of comparing modelled with empirical data.

## METHODOLOGY

---

The main aim of this chapter is to discuss the kinds of evidence used in this thesis, the methods of obtaining this evidence, and how such evidence can be used to address the research question, namely to what extent can the theory of neural resonance provide an account for human perception of polyrhythms. In order to address this question we will introduce and reflect on issues related to the identification and collection of modelled and empirical data, and discuss issues related to the drawing of inferences from comparisons of such data. In particular, the aforementioned issues are directly related to the three different methods used in this thesis.

The first method describes the process of running simulations (tests) with the computational model of neural resonance in order to collect modelled data in response to selected polyrhythmic stimuli, while the second method describes the process of obtaining empirical data derived from human tapping behaviour in response to polyrhythms. While some relevant human empirical data can be found from the literature, others need to be obtained by conducting novel experiments. The third method describes the process of comparing modelled data with empirical data. In general, there are two types of comparisons between modelled and empirical data, each one associated with one of the two halves of the research question. The first type of comparison between modelled and empirical data considers the modes of periodic tapping along with a polyrhythmic stimulus. In this case, existing empirical data are borrowed from the study of Handel and

Oshinsky (1981) and modelled data are obtained from running simulations with the neural resonance model, in particular, by referring to the frequency response pattern of the model. This type of comparison is the theme of Chapter 6 and it aims to answer the first half of the research question, which is: To what extent can the theory of neural resonance provide an account for existing empirical data on how humans perceive polyrhythms? The second type of comparison considers the time it takes for people to start tapping along with a given polyrhythmic stimulus. In this case, the model makes a prediction based on a hypothesis suggested by Large (2000), which is described in detail in Subection 5.2.1 of this chapter, and experiments are conducted in order to obtain empirical data regarding the time it takes for people to start tapping along with a polyrhythmic stimulus. This type of comparison between predicted and empirically observed time to start tapping is the theme of Chapter 8 and it aims to answer the second half of the research question, which is: Can we make novel predictions about human behaviour in polyrhythm perception based on the theory of neural resonance and subsequently test these predictions? The studies in Chapter 6 and 8 can be thought of as validation studies for testing the theory of neural resonance against empirical observations in polyrhythm perception. For that reason, the chapter begins by briefly reflecting on the methodological approach of validation studies.

### 5.1 THE METHODOLOGY OF VALIDATION

Validation encapsulates one of the fundamental methodological rules in scientific discovery, which is the need for the testing of theories against experience by observation. In general, validation can be thought of as the process of exposing a theory to empirical scrutiny. The nature of inferences

deriving from the comparison between real world and simulated data may among other options lead to falsifying or corroborating a theory. According to Popper (1959), corroboration refers to the process, whereby a theory withstands detailed and severe tests and is not superseded by another theory in the course of scientific progress, and as a result we may say that it has 'proved its mettle' or that it is corroborated by past experience. In this thesis the results of the validation studies presented in Chapter 6 and 8 suggest that the theory of neural resonance is corroborated further.

In the remainder of this chapter, we introduce in detail the methods involved in this thesis. In particular, we begin by describing the method of acquiring modelled data by running simulations with the model of neural resonance, and we continue with the method of obtaining empirical data by conducting experiments. Finally, we describe the method of comparison between modelled and empirical data, whereby responses of the neural resonance model in the presence of polyrhythmic stimuli are compared with empirical data of human tapping behaviour in polyrhythm perception.

## 5.2 SIMULATIONS (TESTS)

In order to explore the extent to which the neural resonance theory provides an account for empirical data on human perception of polyrhythms we need to choose appropriate stimuli to feed into the neural resonance model, and then collect data on how the model behaves in response to these stimuli. This behavior can then be compared with human tapping behavior in response to the same stimulus. Recall that the model of neural resonance is a bank of oscillators tuned to a range of frequencies from low to high, reflecting the hypothesized existence of a number of neural oscillators in the brain that can potentially respond to the presence of



rhythmical stimuli. The computational model of the neural resonance, as this was described in Chapter 4, has been developed by Edward Large and his colleagues and it is implemented as a MATLAB toolbox for research in nonlinear resonance using gradient frequency neural oscillator networks. This model is not generally available, but we were able to obtain it, with full documentation and source code for the purpose of this study. Below we consider the following components related to setting up the neural resonance model for the purposes of running the simulations:

- a) Input encoding
- b) Frequency range of the bank of oscillators
- c) Number of oscillators
- d) Duration of simulation/stimulation
- e) Connectivity of oscillators and number of networks

### 5.2.1 *Input encoding*

In order to run simulations using the neural resonance model there is a technical choice to be made regarding encoding a polyrhythmic stimulus. There are three possible ways: functions, audio files sampling the stimulus and MIDI files representing times of events. For the series of simulations related to the discussion in Chapter 6 and Chapter 8 we have utilised the bump function (a pulse generating function) to represent the polyrhythmic stimulus. However, we have conducted additional tests using both of the alternative methods mentioned above with no major effect in altering the results presented here (see Chapter 7).

The choice of polyrhythmic stimuli presented to the model in this thesis takes advantage of their link to existing empirical data related to

polyrhythm perception, allowing direct comparison with the output of the simulations. The study of Handel and Oshinsky (1981) is the main reference for addressing the aforementioned issues.

In the Handel and Oshinsky study the polyrhythmic stimulus presented to people consists of two individual pulse trains, each of which is delivered through a different loudspeaker placed in front of the participants. There are five different types of polyrhythms presented at ten different rates. All polyrhythms are combinations of two pulse trains exclusively. The options for each pulse train are a 2-pulse, a 3-pulse, a 4-pulse or a 5-pulse train. In Chapter 6 an extensive description of the polyrhythmic stimulus is given. The physical characteristics of the stimulus, such as the amplitude and the pitch of each pulse train, and the presentation tempo, are experimental variables that have been found to influence the way people perceive rhythm (Handel and Oshinsky, 1981). For example, in some of their experiments where the focus is on examining the influence of the tempo in polyrhythm perception ( i.e., both the amplitude and the pitch for each pulse train are kept the same), in the case of a really fast repetition rate, e.g., 400 msec, of a 4:3 polyrhythmic stimulus, people choose to tap along with the coincidence of the two pulse trains every 400 msec. For slower repetition rates e.g., 1200 msec, people choose to tap along with either of the pulse trains. For the purposes of this thesis, there is a particular advantage of having a pool of existing empirical data that focus exclusively on individual physical characteristics of the stimulus that may influence the way humans perceive rhythm in a polyrhythm. The advantage relates to the fact that the current version of the neural resonance model is limited in taking into account the full range of physical characteristics of a stimulus such as pitch information. Physical characteristics such as pitch and timbre that has been found to influence rhythm perception, e.g., human preference for attending

to a low-pitch pulse train than a high-pitched (Handel and Oshinsky, 1981). From that perspective we can conduct tests with the neural resonance model that fully resemble existing experimental set ups, i.e., we can feed into the model the same polyrhythmic stimulus that people were listening to and subsequently, to compare human tapping behaviour with the output of the neural resonance model.

### 5.2.2 *Frequency range of the bank of oscillators*

The range of natural frequencies of the bank of oscillators is defined based on the following two aspects of human rhythm perception. The first aspect is related to the sensorimotor-synchronisation (SMS) range of human tapping behaviour. This range helps us define the highest and the lowest frequency of the oscillators in the bank in a way that reflects the highest and lowest tapping frequencies empirically observed in sensorimotor synchronisation tasks (Repp, 2005). The identification of those limits is based on reviewing the literature (London, 2004; Snyder and Large, 2004; Pressing et al., 1996; Repp, 2005). The second aspect is related to the observed tapping frequencies in polyrhythm perception reported by Handel and Oshinsky (1981), which inform us about the granularity of the range.

The human perception of beat and meter implies that events can be perceived as being both discrete and periodic at the same time, that is events are perceived distinctively in a regular manner. The lower inter-onset interval (IOI) necessary for perceiving two events as being distinct is 12.5 msec (Snyder and Large, 2004). With regard to perceiving (or tracking) a series of events as being periodic, London (2004) suggests that the 100 msec interval corresponds to the shortest interval we can hear or perform as an

element of a rhythmic figure. This limit is well documented in the literature as the lower limit for perceiving potential periodicity in successive events (London, 2004; Snyder and Large, 2004; Pressing et al., 1996; Repp, 2005).

Regarding the upper limit for perceiving periodic events there are several suggestions about how long it can be, and as Repp (2005) notes, it is a less sharply defined limit. For example, London (2004, p. 27) suggests that the longest interval (upper limit) we can "hear or perform as an element of rhythmic figure is around 5 to 6 secs, a limit set by our capacities to hierarchically integrate successive events into a stable pattern". Repp (2005) quotes Fraisse's upper limit of 1800 msec, and in a more recent study (Repp, 2008) he points out to the difficulty in providing evidence towards a sharply defined upper limit of sensorimotor synchronisation tasks (SMS). In this study, we are mainly interested in capturing tapping frequencies that are observed in Handel and Oshinsky's study and therefore, we consider the upper limit with regard to these observations. As previously mentioned, we focus on tapping behaviours observed in a case where the sole experimental variable is the tempo of the polyrhythm. In such a situation, the concrete behaviour with which it is appropriate to associate an upper limit on periodicity is tapping along with the unit pattern of the polyrhythm for a repetition rate of 1 second or 1000 msec.

As a result we can set the natural frequencies of the oscillators of the neural resonance model in a way that is arbitrary, but nonetheless informed by the following considerations:

- the limits observed in SMS studies (as noted above),
- the tapping frequencies observed in Handel and Oshinsky's study,
- potential tapping frequencies related to plausibly perceivable repetitive patterns in a polyrhythmic stimulus.

For clarity of presentation it helps to express the sensorimotor synchronisation (SMS) range in terms of frequencies, trivially by making use of the  $F=1/T$  formula, where  $F$  is the frequency and  $T$  is the tapping period in seconds. For example, tapping once every 100 msec implies a tapping frequency of 10 Hz, while tapping once every 2000 msec implies a frequency of 0.5 Hz.

Given our focus on polyrhythms as stimuli, Table 5.1 below illustrates an example of potential tapping frequencies associated with a 4:3 polyrhythmic stimulus for the range of presentation rates studied by Handel and Oshinsky (1981). The repetition rate of the polyrhythmic pattern is given in terms of seconds in the first row, and it has been transformed in frequency terms in the subsequent rows (rounded to 2 d.p where necessary).

Table 5.1: Potential periodic human tapping behaviours along with a 4:3 polyrhythm expressed in frequencies for a series of repetition rates.

Repetition Rate in sec	2.4	2.0	1.8	1.6	1.4	1.2	1.0	0.8	0.6	0.4
Pattern's Frequency	0.42 Hz	0.50 Hz	0.56 Hz	0.63 Hz	0.71 Hz	0.83 Hz	1.00 Hz	1.25 Hz	1.67 Hz	2.50 Hz
Fundamental of 4-pulse train	1.67 Hz	2.00 Hz	2.22 Hz	2.50 Hz	2.86 Hz	3.33 Hz	4.00 Hz	5.00 Hz	6.67 Hz	10.00 Hz
1st sub-harmonic of 4-pulse	0.83 Hz	1.00 Hz	1.11 Hz	1.25 Hz	1.43 Hz	1.67 Hz	2.00 Hz	2.50 Hz	3.33 Hz	5.00 Hz
Fundamental of 3-pulse train	1.25 Hz	1.50 Hz	1.67 Hz	1.88 Hz	2.14 Hz	2.50 Hz	3.00 Hz	3.75 Hz	5.00 Hz	7.50 Hz
1st sub-harmonic of 3-pulse	0.63 Hz	0.75 Hz	0.83 Hz	0.94 Hz	1.07 Hz	1.25 Hz	1.50 Hz	1.88 Hz	2.50 Hz	3.75 Hz

Once the frequency range to be incorporated into the neural resonance model is decided in principle, e.g., 0.42 Hz to 10 Hz according to the minimum and maximum values of Table 5.1, this decision needs to be executed within the constraints of the controls of the software that implements the neural resonance model. In practice, the current version of the software requires the selection of a central frequency and an even number of octaves. The octaves are added symmetrically on either side of the central frequency. So for example if we want a frequency range that covers a tapping frequency range between 0.42 Hz and 10 Hz we can set the central frequency of the bank at 2Hz and then we could add three octaves on either side, i.e., six octaves in total. In this case, the frequency range will be from 0.25 Hz to 16 Hz.

### 5.2.3 *Number of oscillators*

One can define the granularity of the frequency range by making sure that for each empirically observed or potential tapping frequency there is an oscillator with a closely related natural frequency. In fact, maximum resonance occurs for particular ratios such as 1:1, 2:1, 3:2, etc, thus, we have to make sure that the oscillators are packed tightly enough in frequency space in order to match the frequencies defined by the above ratios. For example, for a 4:3 polyrhythm repeating once per second, the frequency of the 4-pulse train is approximately 6.66 Hz. As a result, the bank of oscillators should contain oscillators with frequencies such as 6.66 Hz,  $2 * 6.66\text{Hz}$ ,  $(3/2) * 6.66\text{Hz}$  and so on. A large number of oscillators per octave such as 128 provide a sufficient level of granularity. We have also tried 256 oscillators per octave with no observed difference in the way the neural resonance model responds. As a result the total number of oscillators is

768 inasmuch as there are 6 octaves and each octave accommodates 128 oscillators.

#### 5.2.4 *Duration of simulation*

The duration of simulation of the model should be long enough in order for the oscillators to reach a steady state. One way that steady states can be determined is by obtaining spectrograms covering the simulation period and identifying a point in time at which the amplitude responses of the oscillators appear to be stable. Figure 5.1 below illustrates a simulation period of 7 seconds using a 4:3 polyrhythm which is repeated once every second. The y-axis represents the frequency range of the oscillators, i.e., 0.25 Hz to 16 Hz. The redness indicates amplitude levels and in effect shows which oscillators resonate the most in the presence of this particular polyrhythmic stimulus. For example the oscillators with 3Hz and 4Hz natural frequencies exhibit the highest amplitude (at least for the illustrated time period), followed by oscillators, which are harmonically and subharmonically related to the 3 Hz and 4 Hz frequencies. The x-axis represents time, which manifests the evolution of the dynamics of the system over the duration of the simulation. This graph allows us to illustrate the importance of allowing sufficient time for the system to reach a steady state and also the importance of averaging the activity only after a steady state has been reached. For example, if we had only considered the first two seconds of stimulation and averaged over 100% of that duration we would have observed only dynamics over the transient phase of the system. In that case important information about the sub-harmonic responses of the system with regard to the fundamental frequencies of the stimulus may have been lost, which is a crucial piece of information for the arguments

made in this thesis. From that perspective, a duration of 12 seconds for stimulating the neural resonance model with a polyrhythmic stimulus is sufficient to obtain a frequency response pattern whereby the resonant oscillators are in steady state.

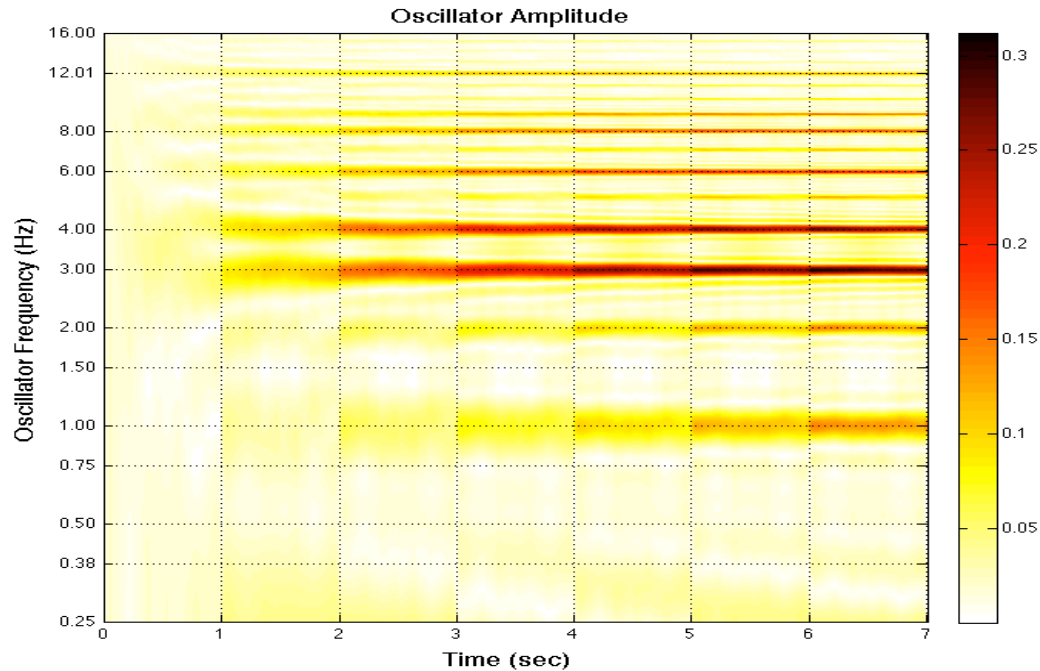


Figure 5.1: Amplitude response of oscillators to a 4:3 polyrhythmic stimulus. The polyrhythmic pattern repeats once every second over a period of 7 seconds.

An interesting aspect with regard to the time it takes for the oscillators to reach a steady state is the potential association of that time with the time it takes for people to start tapping along with a polyrhythm. A hypothesis has been suggested and tested by Large (2000) whereby the time needed for oscillators to reach a steady state squares well with the time it takes for people to start tapping along with a rhythmic stimulus, and as a result it is tested in this thesis with respect to polyrhythms. In the Handel and Oshinsky study the polyrhythmic stimuli were initially presented for 15 seconds, followed by a 5 seconds silent pause after which the participants were asked to start tapping. However, the exact time it



takes before participants start tapping is not documented in that paper and further research did not result in identifying any other paper that explores this aspect of human behaviour. The hypothesis formulated by Large (2000) in combination to the fact that Handel and Oshinsky (1981) did not consider the time it takes for people to start tapping along with a polyrhythm as an experimental variable opens up an opportunity for making new predictions about human tapping behaviour based on the model's outputs which can be tested by conducting novel experiments with people tapping along with polyrhythmic stimuli (Chapter 8).

#### 5.2.5 *Connectivity and number of networks*

In principle, we can have the oscillators of the bank interconnected to each other in order to form a network of oscillators, which is believed to be a more accurate representation of the underlying connectedness of neural populations in the auditory nervous system (Hoppensteadt and Izhikevich, 1996a). However, in this particular series of simulations, partly for reasons of simplicity, we examine the individual responses of each of the oscillators to the external polyrhythmic stimulus, i.e., there is no internal coupling among the oscillators. This chosen arrangement "focuses the response of the network to the external input" (Large et al., 2010, p. 4). As a further choice in setting up the model, we could choose to have more than one bank of oscillators. One argument in favour of assuming more than one bank of interconnected oscillators (network of oscillators) is related to the fact that processing of any auditory stimulus takes place in more than one area in the brain, thus utilising more than one network would be more plausible. As Velasco suggests (personal communication), two linked networks might be small patches of tissue in the same local brain area, or they could be

two areas that are far away from each other, for instance, primary auditory cortex and supplementary motor area (SMA).

Taking into account the above considerations, in the conducted simulations we have employed a single network of non-interconnected oscillators, which is the simplest option for starting to test the neural resonance model. In effect, the simplified version of the model puts the focus on its non-linear resonance feature, i.e., nonlinear coupling between stimulus and oscillators. Further investigations involving both interconnections and more than one network are reserved for future work.

#### 5.2.6 *Numerical solvers and nominal values of parameters*

The output of the simulations is based on solving numerically a series of  $128 * 6 = 768$  differential equations, each one describing a neural oscillator with a unique natural frequency within the 0.25 Hz to 16 Hz range, in the presence of a particular polyrhythm, i.e., type and presentation rate, over a particular duration (simulation time), and a set of nominal numerical values of the parameters that define the dynamic mode of the oscillator.

The nominal values are defined on the basis of certain criteria such as the fact that we are interested in the dynamics of an oscillator near an Andronov-Hopf bifurcation point. However, the selection of certain numerical solvers and numerical values for running the simulations raises questions regarding the stability and consistency of the model in producing certain results. Additionally, the derivation of Equation 7 in Chapter 4 is a canonical representation of a neural oscillator as a dynamic system and as a result it is more abstracted at least in comparison to the Wilson and Cowan (1972) model with its extended range of physiological parameters. The above concerns can be thought of as sources of uncertainty in the

modelling process in general and in the observed outputs of the neural resonance model in particular. For that reason the notion of validation can be extended to include additional techniques, such as parameter sensitivity analysis and cross-model comparison that can help with limiting the degree of uncertainty. The above issues are addressed in Chapter 7 - Model Analysis. Additionally, detailed discussion on the nominal values of the parameters is also reserved for Chapter 7.

### 5.3 EXPERIMENTAL METHOD

This section is devoted in describing the experimental method of this thesis. Briefly, the method refers to conducting experiments with people listening to polyrhythms and monitoring tapping behaviours. The reason for conducting novel experiments with people listening to polyrhythms has been briefly touched upon in Subsection 5.2.4 of this chapter. For example, in discussing how to define the duration of the simulation an observation was made whereby the time it takes for people to start tapping along with a polyrhythm can be considered as an experimental variable itself. Also, results from the first study, which are presented in Chapter 6, show that the model predicts a wider range of tapping behaviours compared to the one reported by Handel and Oshinsky (1981), therefore an additional reason for conducting experiments with people has been identified. Lastly, conducting experiments with people listening to polyrhythms is a good opportunity to test for reproducibility of the data provided by Handel and Oshinsky.

### 5.3.1 *Participants*

There were a total of 37 participants. They were recruited across faculties at the Open University, UK, and they represented a wide range of musical training and ages. No musical experience or expertise was required. The only requirement to participate in the study was to have a general interest or habit in tapping to music e.g., tapping on the steering wheel while listening to music in the car. The musical background was recorded by asking the participants to fill in a personal data form prior to the experiment. No distinction related to age, sex or musical background was made in the analysis presented in Chapter 8, however we have collected such information. The study was approved by the Open University's Human Research Ethics Committee (HREC) and covered by the Open University's insurance indemnity scheme.

### 5.3.2 *Polyrhythmic stimulus and presentation rates*

The experimental set up is similar to the one described by Handel and Oshinsky (1981). Five types of polyrhythms at ten different rates formed the 50 conditions that were presented to people. The types of polyrhythms are 3:2, 4:3, 5:2, 5:3 and 5:4 and the presentation rates in terms of seconds/pattern repetition are 2.4, 2.0, 1.8, 1.6, 1.4, 1.2, 1.0, 0.8, 0.6, and 0.4 sec. Each pulse train is conveyed by a metronomic click of 30 msec duration, thus the two pulse trains sound identical.

### 5.3.3 *Experimental task*

The participants were asked to listen to the presented polyrhythm and start tapping the perceived beat as soon as they felt it. The task of tapping the beat was distinguished from shadowing or copying the pattern, i.e., participants were asked to tap in a steady pulse-like manner. We have emphasised to the participants to start tapping as soon as they feel the beat in order to be as close as possible to the time it takes for them to start perceiving a beat. Also, they were instructed to retain a single beat - once it was chosen - until the end of the trial. In situations where an odd pulse is presented and people feel like tapping along with every other element of the odd pulse, it consequently forces them into a tapping mode that alternates between in phase and out of phase locking with the unit pattern, which is likely to make them feel like changing beat reference. Additionally, participants were asked to start tapping using their finger by pressing the spacebar on a QWERTY keyboard, and to restrain from start tapping with their foot and then copy this with their finger. On each trial the participants were responsible for initiating the presentation of the polyrhythmic stimulus on demand - with a small delay to provide them time to get ready. Participants were asked to tap for approximately 7-10 times before they could stop the current trial and move on to the next one, or to keep tapping until they feel that their perception of beat was settled. When the polyrhythm is first heard a timer begins. Each time the participant presses the spacebar the current value of the timer in msec is recorded. At the end of the trial the ID of the polyrhythm together with the timing data described above constitute the data associated with one trial or condition. The inter-tap-time in combination with the ID, provide sufficient information to derive which beat was chosen by a participant. For example,

consider one condition where a 4:3 polyrhythm is presented at the rate of one repetition per second. Then if the inter-tap-time data is around 250 msec it means that the participant has chosen to tap along with the 4-pulse train. Also, the first tap is taken to estimate the time it takes for participants to perceive a beat in a given polyrhythm. The same process was repeated for all 50 conditions, and in combination with the 37 participants a pool of 1850 trials was obtained. For the above experimental design we used the programming language Max/MSP to implement a patch that would carry out the task.

#### 5.3.4 *Stimulus presentation*

The experimental session took place in a room with just the experimenter and participant present. Participants were seated in front of two adjacent speakers, each of which presented one pulse train. All patterns were presented at a comfortable listening level. The order of presentation of each condition was random, as was the presentation of each pulse from the speakers.

#### 5.3.5 *Method of comparing modelled with empirical data*

In the last part of this chapter, the nature of comparison between modelled and empirical data is discussed. In particular, there are two aspects of human tapping behaviour in polyrhythm perception that are compared with the responses of the neural resonance model in the presence of polyrhythmic stimuli. The first aspect of human tapping behaviour in polyrhythm perception is the range of periodic tapping modes, i.e., tapping frequencies, that people exhibit when confronted by a polyrhythmic

stimulus. The second aspect is the time it takes for people to start tapping along with a polyrhythmic stimulus.

In the first case the comparison between human tapping modes and responses of the neural resonance model in the presence of a commonly shared polyrhythmic stimulus is conducted on the following basis. The human tapping modes are expressed in terms of tapping frequencies (see for example Table 5.1 - Subsection 5.2.2) for different types of polyrhythms and different presentation rates. Next, the neural resonance model is stimulated with a polyrhythmic stimulus from the same range as the one presented to people and a frequency response pattern emerges. The frequency response pattern is the average amplitude response of each oscillator in the presence of a polyrhythmic stimulus over the last 20% of the total duration of the simulation. The reason for averaging over the last 20% percent is due to the need of observing the dynamics of the oscillators once a steady state is reached, i.e., ignoring the transient stage, as has already been described. The comparison between empirical and modelled data is then based on identifying the extent at which the resonances of the frequency response pattern of the model match human tapping frequencies. An example of such comparison was provided in Chapter 4, whereby the frequency response pattern of the neural resonance model in the presence of a series of metronomic clicks has been matched with the way humans perform various levels of the metrical structure such as eighth, quarter, half and sixteenth notes. Chapter 6 is devoted in the above type of comparison for all types and tempi of polyrhythms presented in the Handel and Oshinsky study. In particular, the types of polyrhythms that are fed into the neural resonance model are 3:2, 4:3, 2:5, 3:5, 4:5 and the repetition rates in sec are 2.4, 2.0, 1.8, 1.6, 1.4, 1.2, 1.0, 0.8, 0.6 and 0.4. The

obtained frequency response patterns of the model are then compared with the human tapping frequencies reported by Handel and Oshinsky (1981).

In the second case, the comparison of human behaviour and model responses is made by comparing the time it takes for people to start tapping along with a polyrhythm and the time it takes for an oscillator to reach relaxation time. This comparison is based on a hypothesis made by Large (2000) who observed that the time it takes for humans to start tapping along with a stimulus squares well with the relaxation time of an oscillator. Chapter 8 is fully devoted in presenting the results of the above comparison. Empirical data for this comparison are obtained by conducting novel experiments following the method described in Section 5.2.2. The modelled data are obtained from the amplitude response graphs of the bank of oscillators (see Figure 5.1) for different types of polyrhythms and presentation rates.

#### 5.3.6 *Justification of selected method of comparison*

We should note that the neural resonance model simulates populations of neurons and their dynamics in the presence of a rhythmical stimulus over time. From that perspective, the empirical data should ideally be collected from monitoring electrical activity of the brain over time by using brain-imaging techniques, and in particular, imaging of the brain while it is confronted by a polyrhythmic stimulus. However, obtaining this kind of empirical data requires identification and spatial mapping of each single neural oscillator independently, which is beyond the reach of this thesis. Identifying individual oscillators that behave in the way the neural resonance theory hypothesises would have provided strong



empirical evidence in support of the theory, but this type of evidence it is currently impractical to obtain.

In fact, most theories that attempt to explain human behaviours do not involve descriptions at the level of neuronal activity. Instead, the approach is to explain how brain dynamics might implement abstract computational mechanisms required by cognitive theories, e.g., (Jackendoff, 2003). The theory of neural resonance on the other hand can be thought of as a behaviour-level theory according to which musical behaviour may not require postulation of abstract computational mechanisms, but may be explainable directly in terms of neurodynamics (Large et al., 2010). From that perspective, a comparison that involves simulated brain dynamics of the neural resonance model with empirical data related to tapping behaviours seems to be reasonable. Indeed, a study by Pollok, Gross, Müller, Aschersleben, and Schnitzler (2005) has shown that when people are asked to tap along with a periodic stimulus, activity in the brain that shares the same frequency with the tapping frequency can be observed. Additionally, brain activity with frequency that corresponds to the second and third harmonic of the tapping frequency has also been identified. Lastly, in Chapter 4, a number of existing validation studies of the neural resonance model were described in which the standard technique is to compare modelled brain dynamics with human tapping behaviours observed in rhythm perception experiments, which is exactly the same method of comparison as the one utilised in this thesis.

## VALIDATION OF THE MODEL AGAINST EXISTING EMPIRICAL DATA

---

In this chapter we examine the extent to which the behaviour of the neural resonance model in the presence of polyrhythmic stimuli can account for polyrhythm perception as this is manifested by tapping behaviours reported in the empirical study of Handel and Oshinsky (1981). In particular, we express regular tapping behaviours by people in terms of tapping frequencies, which can then be compared with the frequency response pattern of the neural resonance model. The neural resonance model was stimulated with similar polyrhythmic stimuli to those used by Handel and Oshinsky (1981) in their experiments with human subjects. In particular, Handel and Oshinsky used polyrhythmic stimuli comprising of two pulse trains each one perfectly isochronous, i.e., completely regular.

Furthermore, the model was stimulated with the same type of polyrhythms, but this time the pulse trains exhibit temporal fluctuations with respect to perfect isochrony. This is to have a form of stimulus that resembles more the context within actual music is performed and perceived. Two different stimuli were used to stimulate the neural resonance model. In the first one, a musician performs a polyrhythm in the presence of a metronomic click, thus tempo variation is restricted but local deviations from perfect isochrony are expected. In the second case, the polyrhythm is performed in the absence of a metronomic click (after four initial clicks), which allows for tempo variation over time as well as local deviations from

perfect isochrony. The last two cases aim to test the model with stimuli that resemble the way actual polyrhythmic music is performed.

Lastly, we explore whether the behaviour of the model can provide further insight with regard to the formation of human tapping preferences as a function of absolute tempo (Handel and Oshinsky, 1981). In particular, we examine whether different tempi of a particular polyrhythm influence the way the model behaves and whether an assumed difference in the behaviour of the model can be somehow linked to the patterns of preferences empirically observed.

The chapter begins by restating the categories of human tapping behaviours in polyrhythm perception observed in the Handel and Oshinsky study (see Chapter 2), which provides the reference point for making comparisons between empirical and modelled data. Several different polyrhythms over a range of tempi are featured in Handel and Oshinsky's work, and also in the present research. It is useful for purposes of exposition, and in order to avoid confusion, to select one representative polyrhythm to use as a workhorse example. We have chosen the 4:3 polyrhythm with 1 Hz frequency repetition as a good representative for the family of polyrhythms consisting of two pulse trains (3:2, 2:5, 3:5, 4:5). The output of the neural resonance model is then presented and the extent to which human tapping behaviour can be understood on the basis of that output is discussed.

## 6.1 PREPARING TO COMPARE HUMAN WITH MODEL BEHAVIOUR

In Chapter 2 four categories were identified to describe the way people perceive beat and/or meter in polyrhythms. In Handel and Oshinsky's

words (1981, p.2):

"Polyrhythms permit many possible meter organisations: (1) a pulse train can define the meter, achieved by tapping each event in that pulse train; (2) a pattern repetition can define the meter, achieved by tapping each co-occurrence of the 2-pulse trains; and (3) alternative elements of a pulse train can define the meter, achieved, for example by tapping every second element of a 4-pulse train" .

For example, given a 4:3 polyrhythm the general categories identified by Handel and Oshinsky with regard to how people perceive beat and meter in polyrhythms are as follows.

1. Tapping along with the coincidence of the two pulse trains
2. Tapping along with the 4-pulse train
3. Tapping along with every other element of the 4-pulse train
4. Tapping along with the 3-pulse train
5. Tapping along with every other element of the 3-pulse train

As previously noted, in order to facilitate comparison between the tapping behaviours listed above and the frequency response pattern of the neural resonance model in the presence of some rhythmic stimulus it is useful to express the former in frequency terms. Recall that the way the model responds in the presence of a rhythmical stimulus is to resonate at frequencies related to the metrical structure of the stimulus (see Chapter 4, 4.5). Therefore by expressing tapping behaviours in frequency terms we can make a direct comparison between the frequency response pattern of the neural resonance model and the tapping frequencies for a commonly

shared type of polyrhythmic stimulus, e.g., a 4:3 polyrhythm with 1 Hz repetition frequency.

Table 6.1 below expresses in frequency terms the human tapping behaviours listed above for a case where the 4:3 polyrhythmic pattern is repeated once per second<sup>1</sup>.

Table 6.1: Periodic tapping behaviours along with a 4:3 polyrhythmic stimulus expressed in terms of frequencies, when the repetition frequency of the polyrhythmic pattern is 1 Hz.

1. Tapping along with the 4-pulse train.	4 Hz
2. Tapping along with the 3-pulse train.	3 Hz
3. Tapping along with the co-occurrence of the two pulses.	1 Hz
4. Tapping along with every second element of the 4-pulse train.	2 Hz
5. Tapping along with every second element of the 3-pulse train.	1.5 Hz

Before we move on to consider the way the neural resonance model responds to the presence of such a stimulus it is worth noting and briefly analysing the relation of the tapping behaviours to the constituent pulse trains. On the one hand, we should note that in the above example the polyrhythm consists of two pulse trains exclusively, with rhythmic frequencies of 4 Hz and 3 Hz. On the other hand, the final three tapping behaviours noted in Table 6.1, cases 3 to 5, imply a tapping behaviour that is subharmonically related to the fundamental frequencies of the two physically presented pulse trains. This kind of relation represents an important characteristic of human rhythm perception, which is the ability of people to perceptually construct a rhythm that is subharmonically related to a presented rhythm. In fact, this relation indicates the ability of

<sup>1</sup> In Chapter 5, tapping behaviours were interpreted in terms of frequencies for the whole range of polyrhythmic pattern durations in the Handel and Oshinsky study.

people to perceive metrical structure in a rhythmic stimulus, as discussed in detail in Chapter 2 (see Section 2.3). There it was further suggested that any model that aims to simulate rhythm perception should be able to account for this.

## 6.2 SIMULATIONS WITH ISOCHRONOUS POLYRHYTHMIC STIMULI

Our next step is to apply polyrhythmic stimuli to the neural resonance model. As previously noted, for purposes of exposition we have so far focused on a 4:3 polyrhythm presented at 1 Hz, but in fact Handel and Oshinsky used a 5 different types of polyrhythms, and presented these stimuli to human subjects at a range of ten different tempi. A similar range of stimuli and a similar range of tempi to those used by Handel and Oshinsky were applied to the neural resonance. More generally, the polyrhythmic stimuli were structurally regular, i.e., individual pulse trains are isochronous.

### 6.2.1 *Simulation with a 4:3 polyrhythmic stimulus repeating once per second*

The duration of the simulation was decided on the basis of allowing enough time for the system (bank of oscillators) to reach a steady state. In this simulation the period is 12 seconds, which was decided by observing the amplitude response of the bank of oscillators (see Chapter 5, Subsection 5.2.1d). Figure 6.1 is a time series representation of a 4:3 polyrhythmic stimulus repeating once per second composed of two pulses with rhythmic frequencies of 4 Hz and 3 Hz respectively. The stimulus was created using the 'bump' function.

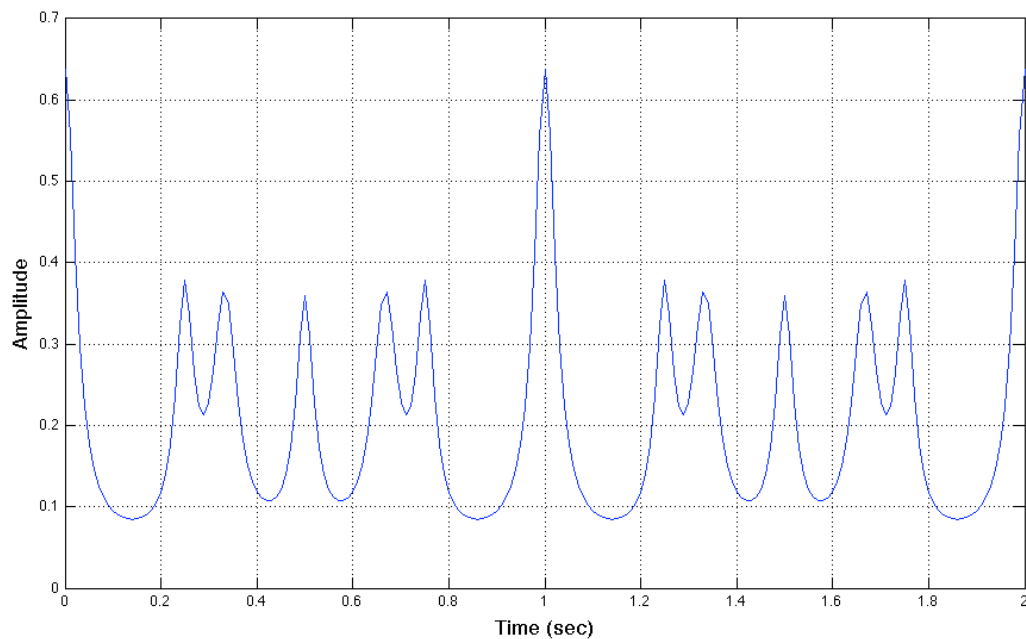


Figure 6.1: Time series representation of the 4:3 polyrhythmic stimulus. The stimulus is a combination of two pulses of 4 Hz and 3 Hz respectively. The graph represents two cycles of completion of the polyrhythmic pattern.

Figure 6.2 below shows the frequency response pattern of the neural resonance model in the presence of the 4:3 polyrhythmic stimulus with 1 Hz presentation rate. The frequency response pattern shown is averaged over the last 20% of the stimulation period, i.e., 2.4 seconds.

Each amplitude peak in the frequency response graph results from the activation of several oscillators (each dot represents one of 768 oscillators). However, in each peak there is generally one particular oscillator that exhibits a maximal amplitude response and we refer to it as the main oscillator. For ease of comparison, in Figure 6.2 some amplitude peaks have been labelled in bold type to indicate that they correspond to the categories of human tapping frequencies described by Handel and Oshinsky. These main oscillators are listed below:

An oscillator with 3 Hz natural frequency

An oscillator with 4 Hz natural frequency

An oscillator with 1 Hz natural frequency

An oscillator with 2 Hz natural frequency

An oscillator with 1.5 Hz natural frequency

In addition to the peaks that correspond to human observed tapping frequencies, there are other peaks. These peaks generally correspond to oscillators with natural frequencies harmonically related to the frequencies of the main oscillators noted above. For example, in Figure 6.2 below we note peaks corresponding to the 2<sup>nd</sup> and 3<sup>rd</sup> harmonics of the 3 Hz and 4 Hz oscillators, and also, peaks corresponding to the 5<sup>th</sup> and 7<sup>th</sup> harmonics of the 1 Hz oscillator. However, these peaks (grey-labelled) do not appear to correspond directly to any human behaviour as these are expressed in terms of categories in the Handel and Oshinsky study. Despite this, some of these harmonic responses, such as 6 Hz and 8 Hz are attainable tapping frequencies by people according to the literature review on the limits of sensorimotor synchronisation studies (see Chapter 5, Subsection 5.2.2).

This kind of observation opens up an opportunity for repeating experiments similar to the Handel and Oshinsky study to refine our understanding about the range of tapping behaviour in polyrhythms further. The results of these empirical studies are presented in Chapter 8.



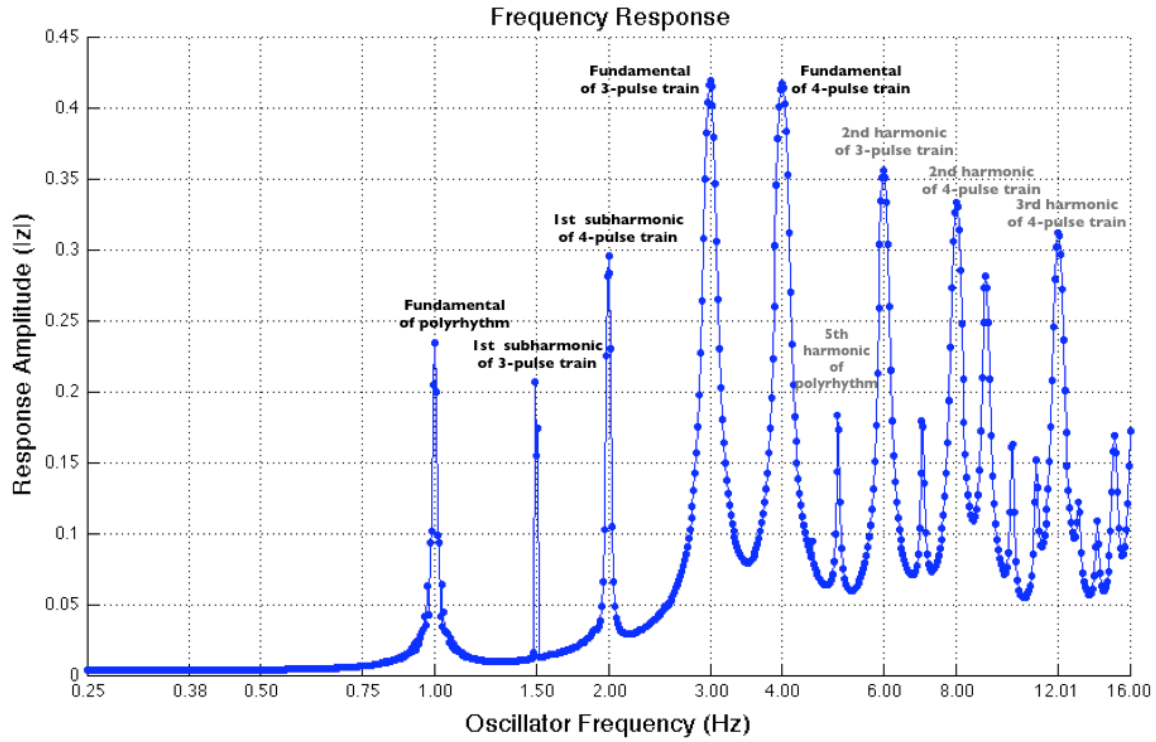


Figure 6.2: Frequency response of the neural resonance model in the presence of a 4:3 polyrhythm. The figure illustrates the average amplitude over the last 20% of the stimulation time. The main oscillators in the black labelled peaks represent oscillators with natural frequencies matching human tapping frequencies (e.g., fundamental of polyrhythm, 1st sub-harmonic of the 3-pulse train, 1<sup>st</sup> sub-harmonic of the 4-pulse train or 2<sup>nd</sup> harmonic of the polyrhythm, fundamental of the 3-pulse and 4-pulse train). The grey-labelled peaks are formed by oscillators with natural frequencies that correspond to harmonics of both pulse trains (e.g., 2<sup>nd</sup> and 3<sup>rd</sup> harmonics of both 4-pulse and 3-pulse train, and harmonics of the polyrhythm's fundamental frequency), and more odd overtones.

### 6.2.2 Discussion on simulation

We have already mentioned that the output of the simulation accounts for all the categories of human tapping frequencies observed in the Handel and Oshinsky study. In particular, with regard to the five tapping behaviours

related to a 4:3 polyrhythmic stimulus (the five cases in Table 6.1), the neural resonance model produces an output that matches all of them.

As mentioned earlier, it is interesting that the output of the neural resonance model includes resonances that are subharmonically related to the fundamental frequencies of the two pulse trains of the polyrhythmic stimulus. There is no pulse train explicitly present in the initial 4:3 polyrhythmic stimulus that matches any of these subharmonic responses, i.e., the polyrhythmic stimulus is solely comprised of two pulses of 4 Hz and 3 Hz. Mathematically, these subharmonic responses by the neural resonance model can be attributed to the expansion of the higher order terms of the canonical formulae described in Chapter 4. In particular, this expansion leads to resonant monomials that correspond to harmonics, subharmonics, and higher order combinations of the input frequencies (Large et al., 2010). It is instructive to compare the frequency response of the neural resonance model with a bank of linear oscillators stimulated with the same 4:3 polyrhythmic stimulus. Figure 6.3 below shows the response of a bank of linear oscillators.

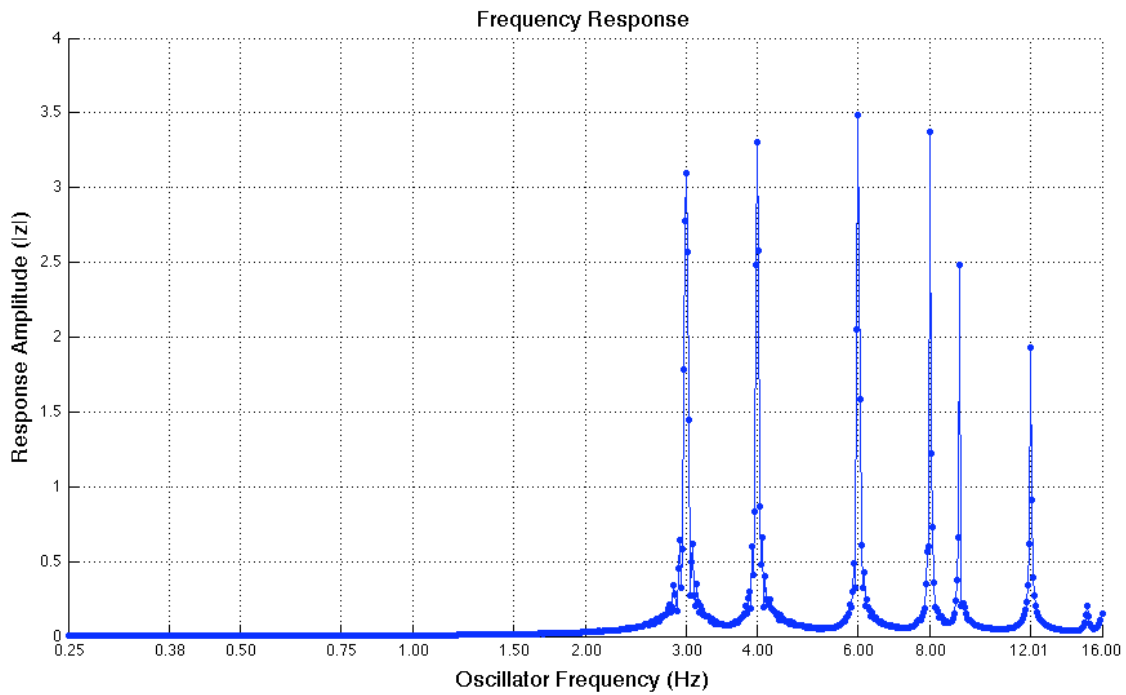


Figure 6.3: Response of a series of linear oscillators in the presence of a 4:3 polyrhythmic stimulus that repeats once per second. The polyrhythm consists of two pulses 4 Hz and 3 Hz.

In comparison to the human tapping behaviours (Table 6.1) and to the output of the neural resonance model (Figure 6.2), it is worth noting that there is no response at all corresponding to any of the subharmonic responses. This bank of linear oscillators is essentially the neural resonance model with its nonlinear parameter ( $\epsilon$ ) set to zero. Thus, the subharmonic responses in the output of the neural resonance model may be attributed to the non-linear coupling between the stimulus and the oscillators.

### 6.2.3 Simulations using the 4:3 polyrhythm at different tempi

By considering a single polyrhythm at a single tempo, we have already been able to demonstrate that the neural resonance model can account for tapping behaviours that are not accounted for by any simple linear

resonator model. However, the available empirical data on human tapping behaviour from Handel and Oshinsky extends much further, since it encompasses different polyrhythms at different tempi. Consequently, the next step is to consider what happens when the neural resonance model is stimulated with the 4:3 polyrhythm, at different tempi.

The following graphs illustrate the frequency response pattern of the neural resonance model in the presence of a 4:3 polyrhythmic stimulus at all the remaining nine repetition frequencies of the Handel and Oshinsky study. The frequency response pattern of the model is in essence the same for all different repetition frequencies, apart from being shifted gradually to the right on the x-axis. As a result, the discussion regarding the frequency response pattern of the model in the presence of a 4:3 polyrhythm repeating once per second applies for all different repetition frequencies presented below. The appropriate duration, for all different tempi, of the application of the stimulus before taking the average frequency response over the last 20% of the stimulation duration was calculated on the basis of allowing sufficient time for the oscillators to reach relaxation (steady state). Given that the lower the frequency of an oscillator the longer it takes for it to relax, polyrhythmic stimuli with slow repetition frequencies will need to stimulate the bank of oscillators for longer periods than stimuli with fast repetition frequencies.

Figure 6.4 shows the frequency response pattern resulting from a 4:3 polyrhythmic pattern stimulus that was repeated once every 2400 msec for approximately 28.8 seconds. The two pulse trains comprising the actual polyrhythmic stimulus presented to the model have rhythmic frequencies of 1.25 Hz (3-pulse train) and 1.67 Hz (4-pulse train). For the other tempi used in presenting the current polyrhythm, information about the tempo and the rhythmic frequencies of the pulse trains comprising the stimulus

are provided exclusively in the captions of the respective figures (Figures 6.4 - 6.12).

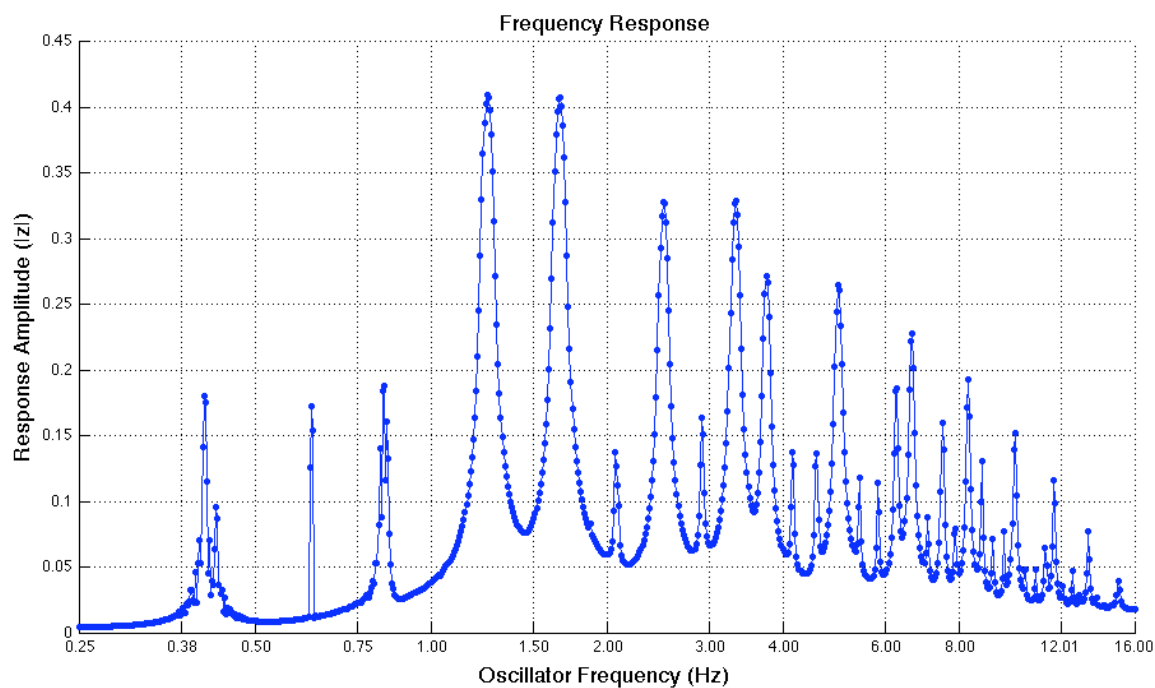


Figure 6.4: Frequency response of the model in the presence of a 4:3 polyrhythm repeating once every 2400 msec. The polyrhythmic stimulus consists of a 1.25 Hz 3-pulse train and a 1.67 Hz 4-pulse train.

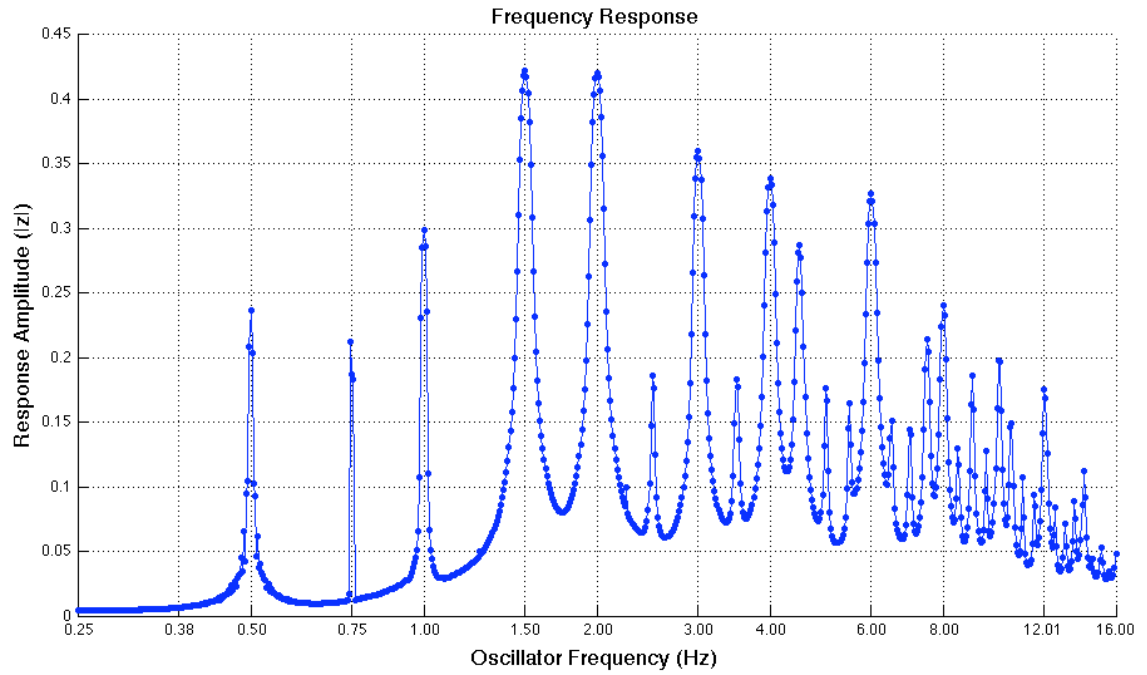


Figure 6.5: Frequency response of the model in the presence of a 4:3 polyrhythm repeating once every 2000 msec. The polyrhythmic stimulus consists of a 1.5 Hz 3-pulse train and a 2 Hz 4-pulse train.

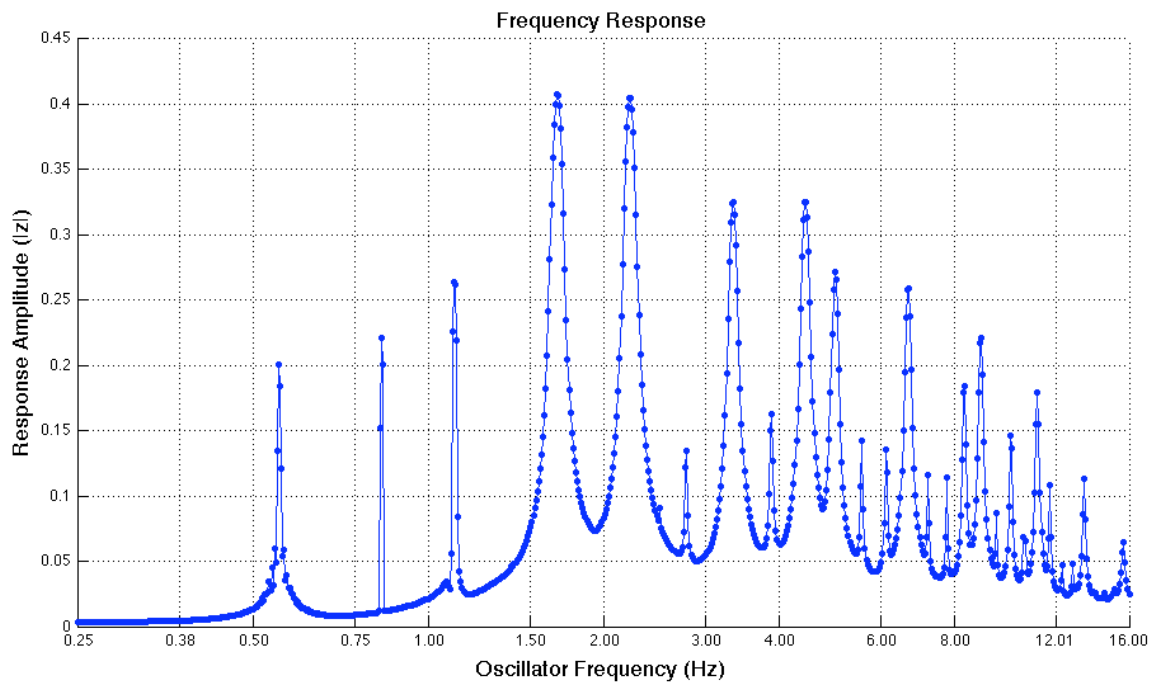


Figure 6.6: Frequency response of the model in the presence of a 4:3 polyrhythm repeating once every 1800 msec. The stimulus consists of a 1.67 Hz 3-pulse train and a 2.22 Hz 4-pulse train.

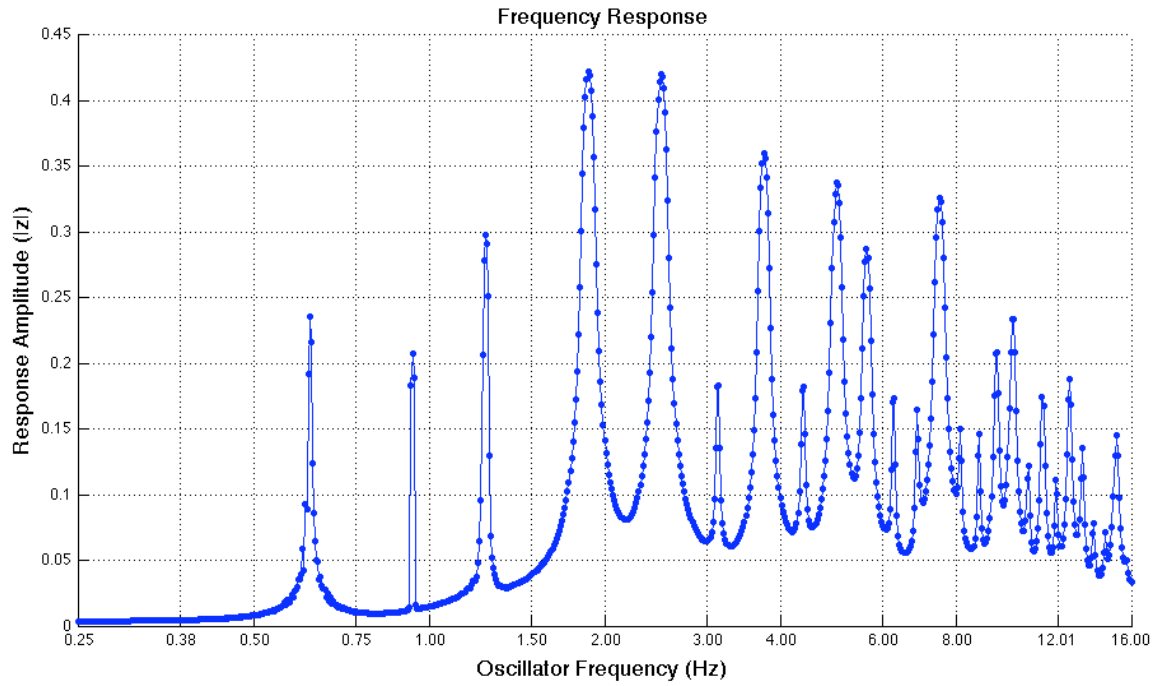


Figure 6.7: Frequency response of the model in the presence of a 4:3 polyrhythm repeating once every 1600 msec. The stimulus consists of a 1.88 Hz 3-pulse train and a 2.5 Hz 4-pulse train.

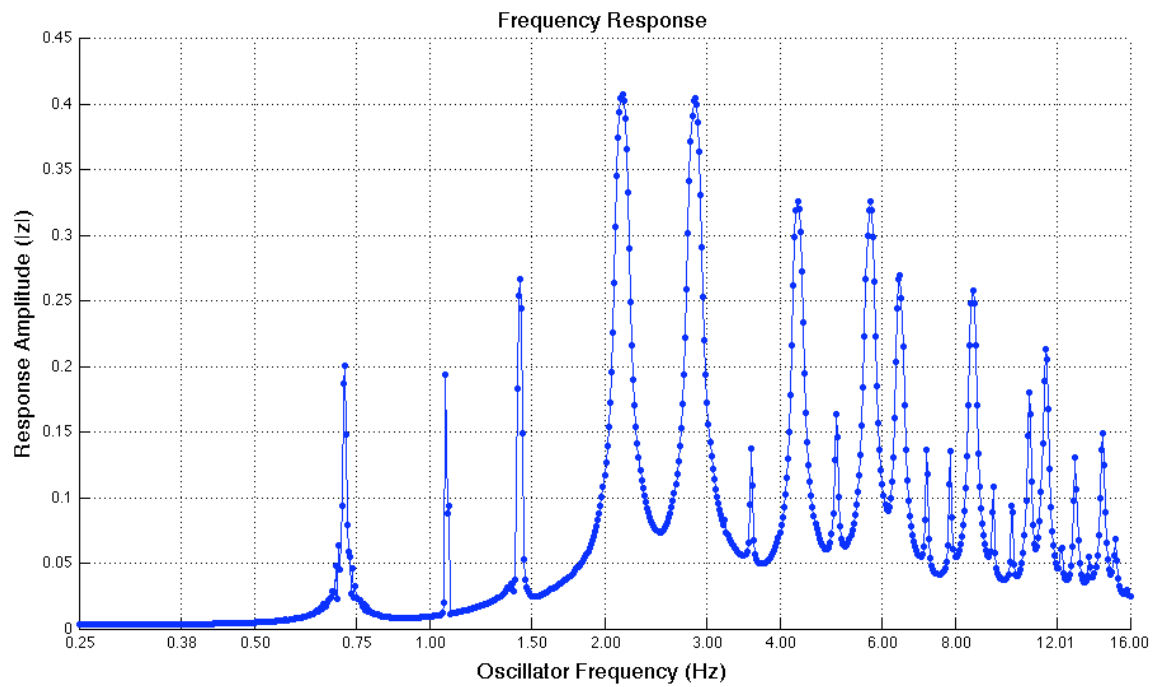


Figure 6.8: Frequency response of the model in the presence of a 4:3 polyrhythm repeating once every 1400 msec. The stimulus consists of a 2.14 Hz 3-pulse train and a 2.86 Hz 4-pulse train.

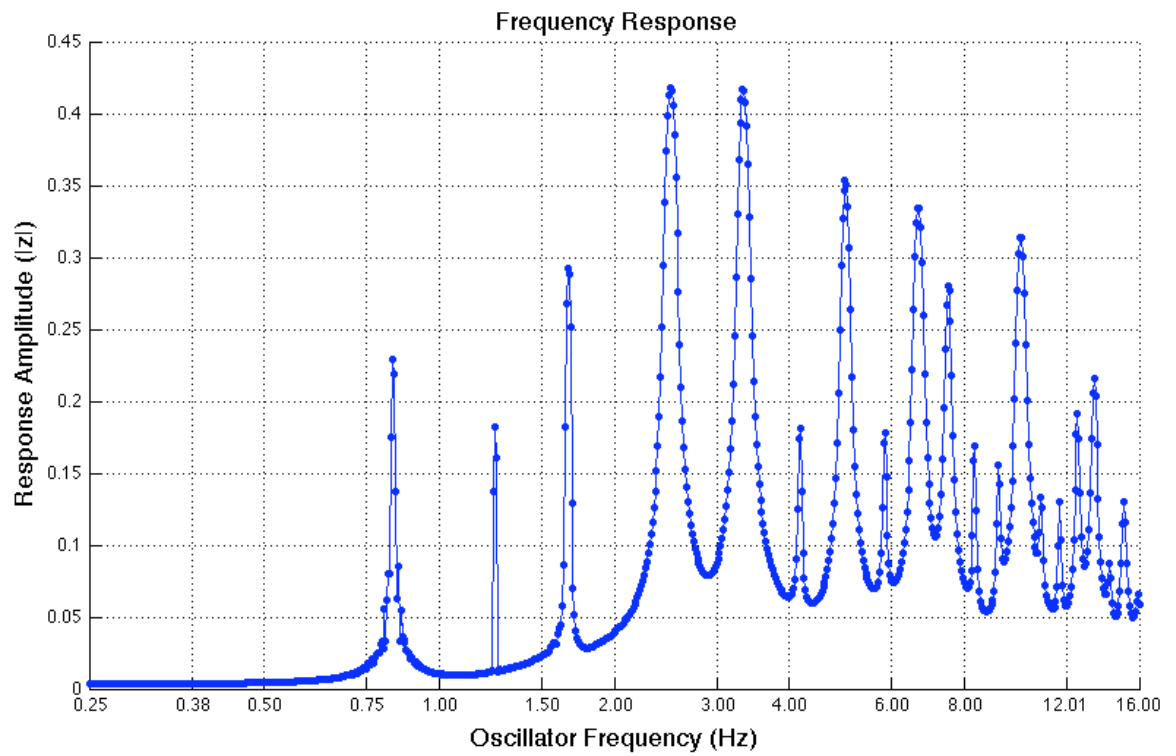


Figure 6.9: Frequency response of the model in the presence of a 4:3 polyrhythm repeating once every 1200 msec. The stimulus consists of a 2.5 Hz 3-pulse train and a 3.33 Hz 4-pulse train.



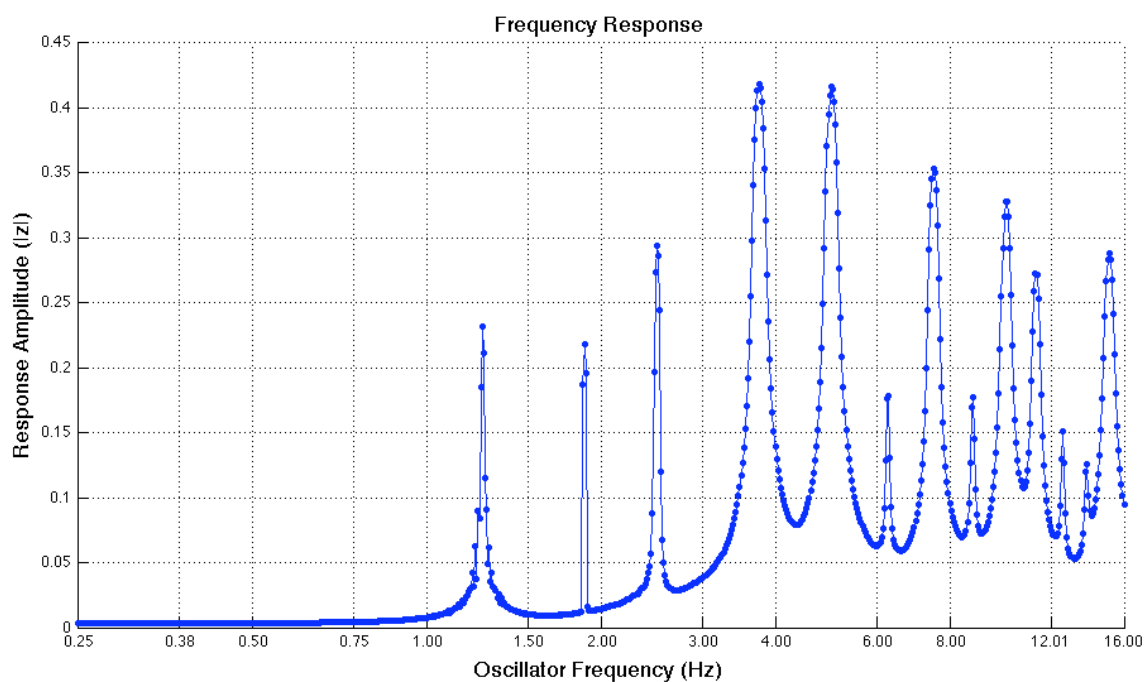


Figure 6.10: Frequency response of the model in the presence of a 4:3 polyrhythm repeating once every 800 msec. The stimulus consists of a 3.75 Hz 3-pulse train and a 5 Hz 4-pulse train.

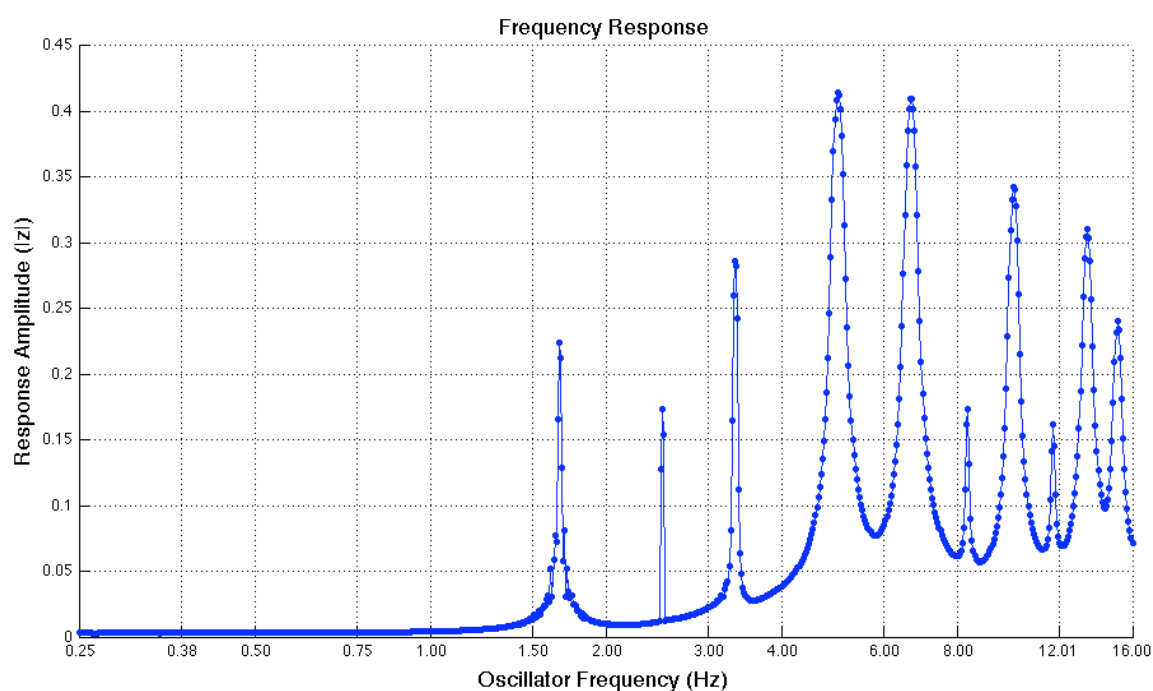


Figure 6.11: Frequency response of the model in the presence of a 4:3 polyrhythm repeating once every 600 msec. The stimulus consists of a 5 Hz 3-pulse train and a 6.67 Hz 4-pulse train.

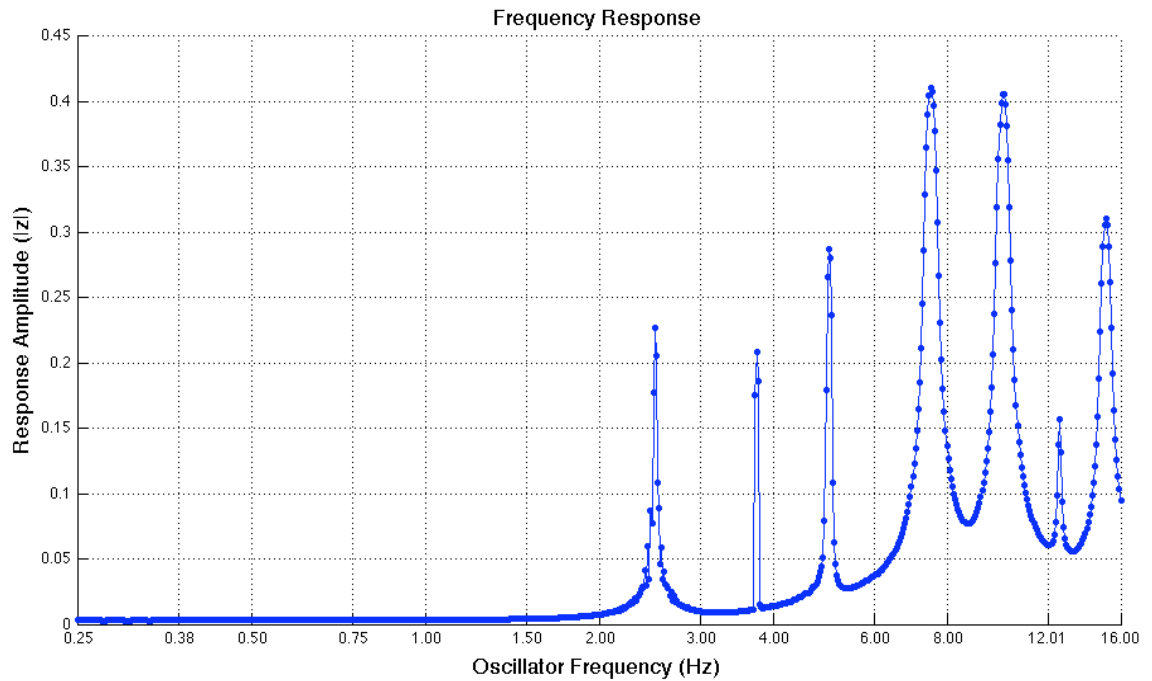


Figure 6.12: Frequency response of the model in the presence of a 4:3 polyrhythm repeating once every 400 msec. The stimulus consists of a 7.5 Hz 3-pulse train and a 10 Hz 4-pulse train.

We have now stimulated the model with a single polyrhythm at a range of tempi. The neural resonance model has been able to account for the range of tapping behaviours empirically observed (Handel and Oshinsky, 1981). Consideration of new tempi has not so far unearthed any new phenomena to be accounted for. In the beginning of the chapter we mentioned that we are interested in examining whether different tempi of a particular polyrhythm influence the way the model behaves and whether an assumed difference in the behaviour of the model can be somehow linked to the patterns of preferences empirically observed by Handel and Oshinsky (1981). Empirical data suggest that a rhythmic figure is perceived differently through transpositions in time. In the case of polyrhythms, the formation of human preferences in perceiving a regular beat along with a polyrhythm and their relative changes through transposition of the polyrhythm in time may be seen as a manifestation of this dynamic

perception. Also, a brain imaging study by Iversen, Repp, and Patel (2009), showed that the interpretation of a rhythm in a particular way was reflected in the magnitude of the amplitude responses of brain activity. In particular, Iversen et al. tested whether voluntary metrical interpretation modulates brain responses. The stimulus consisted of a repeating sequence of two tones followed by a rest (symbolised TT0). The two tones were identical. In different trials, participants were instructed to adopt one of two metrical interpretations of the TT0 phrase by mentally placing the beat on either the first or the second tone. By comparing the event related potentials associated to the presence of each tone they found that amplitude response in the brain was stronger (by 64% on average) when a tone was chosen to represent the beat, thus preference to one specific beat produced enhanced amplitude response in the brain. As a result, a simple hypothesis we wanted to test was to associate the relative preferences for different beats in a polyrhythm with the magnitude of the amplitude response of the corresponding oscillators in the model. For example, the largest (human) preference for a particular way of tapping would be reflected to the strongest amplitude response of a corresponding oscillator in the frequency response pattern of the model.

However, the frequency response pattern remains invariant for different tempi apart for some small amplitude fluctuations in the subharmonic responses with respect to the input frequencies. We have repeated the same tests with a different form of stimulus encoding (square waves) and the frequency response pattern remained completely invariant over different tempi. The reason we have not selected square waves to represent our polyrhythmic stimuli is to avoid a theoretical limitation that follows from the use of square waves for encoding the pulse trains. Namely, the decomposition of a square wave into an infinite set of simple oscillating

functions only contains components of odd integer harmonic frequencies. The bump function, by contrast avoids this limitation by also containing even harmonics of the fundamental frequencies. In Chapter 7 we include superimposed frequency response patterns for the different forms of stimulus encoding. The next step is to consider what happens when the neural resonance model is stimulated with polyrhythms other than 4:3.

In particular, we now consider the frequency response patterns of the neural resonance model in the presence of the remainder of the polyrhythms used by Handel and Oshinsky in their study. These are the 2:3, 2:5, 3:5, and 4:5 polyrhythms. Note that, having observed previously that that tempo does not alter the relative pattern of peaks in the frequency response pattern, in order to simplify the discussion, and without loss of generality for the issues being considered, for purposes of presentation, we will consider the tempo for each type of polyrhythms to be one repetition per second. Thus, in the case of a 3:2 polyrhythm, the constituent pulse trains have rhythmic frequencies 2 Hz and 3 Hz. Consequently, the type of polyrhythm determines the rhythmic frequencies of each pulse train.

#### 6.2.4 *Simulation with a 3:2 polyrhythmic stimulus repeating once per second*

Figure 6.13 below is the frequency response pattern of the neural resonance model in the presence of a 3:2 polyrhythmic stimulus repeating once per second for 12 seconds. The response is the average of the activity over the last 20% of the stimulation period. We can see from the graph that the frequency response pattern is in line with the general framework of empirical findings on how people perceive polyrhythm. With regard to Handel and Oshinsky's categories, the frequency responses at 2 Hz and 3 Hz reflect the first category noted above, i.e., tapping along with each

of the pulse trains. The frequency response at 1 Hz reflects the second category whereby people tap along with the co-occurrence of the two pulse trains. However one may equally interpret this as tapping to every other element of the 2-pulse train. Similarly a frequency response at 1.5 Hz reflects the third category whereby people choose to tap along with every other element of the 3-pulse train.

Interestingly, despite the multiplicity of categories noted in Handel and Oshinsky's framework for classifying how people perceive a beat/meter in polyrhythms, their results for the 3:2 polyrhythm did not include any empirical observations that fell outside the first two categories of their framework (i.e., tapping along with the 2-pulse train, the 3-pulse train, and the co-occurrence of the two pulses). However, this interesting gap may be due to the fact that in their analysis they have combined observations related to the third category with observations related to the first. So for example if a participant chose to tap every other element of the 3-pulse train, in the analysis this behaviour was interpreted as tapping along with the three pulse train since according to Handel and Oshinsky "these responses indicate the salience of the given pulse train" (1981, p. 3).

In the case of the 3:2 polyrhythmic stimulus, as in the case of the 4:3 polyrhythm, the neural resonance model gives frequency responses that cannot be associated with reported tapping behaviours in the empirical study. Of course for frequencies beyond 10 Hz, tapping is quite likely impossible for people, as suggested by data on the upper and lower limits of sensorimotor synchronization tasks (Repp, 2005). However, there is a range of frequency responses, in particular between 4 Hz and 10 Hz, which are potentially attainable by people as tapping speeds. For example the frequency responses at 4 Hz and 6 Hz resemble the 2<sup>nd</sup> harmonics of the fundamental frequencies of the two pulse trains, which could be interpreted

as a tapping behaviour whereby one focuses on either of the pulse trains and taps twice as fast as the rate of the pulse train.

A particularly interesting frequency response occurs at 5 Hz, which resembles the addition of the input frequencies, i.e., 3 Hz plus 2 Hz equals 5 Hz. Although we can imagine people listening to a 3:2 polyrhythm repeating once per second and tapping along with it in a 5 Hz tapping frequency, we can expect that such tapping behaviour is more challenging in comparison to produce tapping behaviours that are harmonically or subharmonically related to the input frequencies.

An alternative explanation, however within a different context, could be the following and it should be taken parenthetically to the main thread of this chapter. We have already discussed that the bank of oscillators of the neural resonance model are originally designed to model actual neural oscillators in the human brain (see Section 4.1). From that perspective, the frequency response pattern of the neural resonance model simulates activity in the brain and in particular neural oscillators resonating to the presence of a polyrhythmic stimulus, so the 5 Hz response can be seen as a component of brain activity. Indeed, Langner (2007) observed that neurons in the inferior colliculus of the gerbil (desert rat) respond at ratios such as 3:2 and 5:2 with respect to the frequency of a pulse-like stimulus. So in the aforementioned example, the 5 Hz response of the neural resonance model can be interpreted as a response along the above lines, i.e., a response that stands in a 5:2 ratio with respect to the 2 Hz input frequency, and it may be understood solely as a neural response.

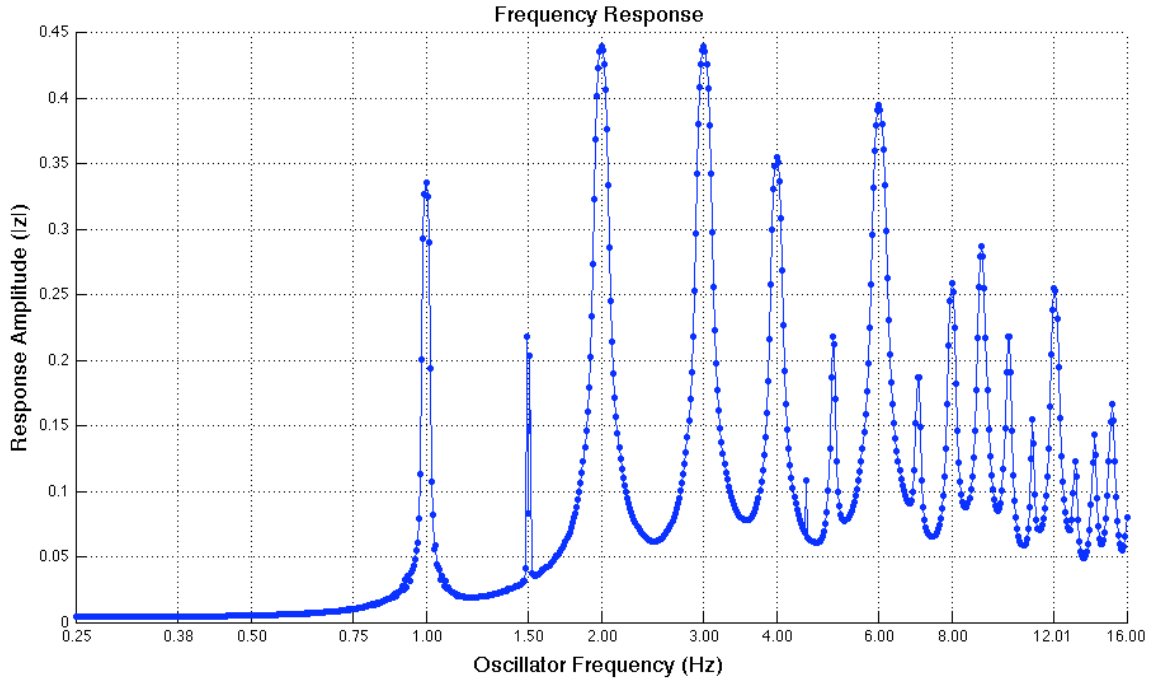


Figure 6.13: Frequency response of the model in the presence of a 3:2 polyrhythm repeating once every 1000 msec. The stimulus consists of a 2 Hz pulse train and a 3 Hz pulse train.

#### 6.2.5 Simulation with a 2:5 polyrhythmic stimulus repeating once per second

In the next simulation the frequency response pattern of the neural resonance model was obtained using a 2:5 polyrhythm as a stimulus. Similarly to the previous simulations the pattern was repeated once per second for reasons explained earlier and consistency of presentation. Figure 6.14 below shows two frequency responses that match the input frequencies, i.e., 2 Hz and 5 Hz, and at least four harmonics of the input frequencies, including the 2nd and 3rd harmonics of the fundamentals, i.e., 4 Hz and 10 Hz (2<sup>nd</sup> harmonics), and 6 Hz and 15 Hz (3<sup>rd</sup> harmonics), and the 4<sup>th</sup> and 6<sup>th</sup> harmonics of the 2 Hz input frequency, i.e., 8 Hz and 12 Hz. Four further responses are subharmonically related to the input frequencies, as follows. The strongest subharmonic response is at 1 Hz, followed by one at 2.5 Hz, and another one at 1.5 Hz. Additionally, there are frequency responses

at 3 Hz and 7 Hz which are of a similar nature to the 5 Hz response in the 3:2 simulation, i.e., implying a response obtained as a result of subtracting and summing the input frequencies respectively. Considering the empirical human data for people tapping along with a 5:2 polyrhythm, Handel and Oshinsky report only a limited range of tapping behaviours compared to the frequency response pattern of the model. In particular, Handel and Oshinsky report that when people are asked to tap along with a 2:5 polyrhythm they choose to tap either along with one of the pulse trains (2-pulse or 5-pulse trains), or along with the coincidence of the pulse trains. These behaviours correspond to a 2 Hz, 5 Hz, and 1 Hz tapping frequencies respectively, all of which are present in the frequency response pattern of the model below.

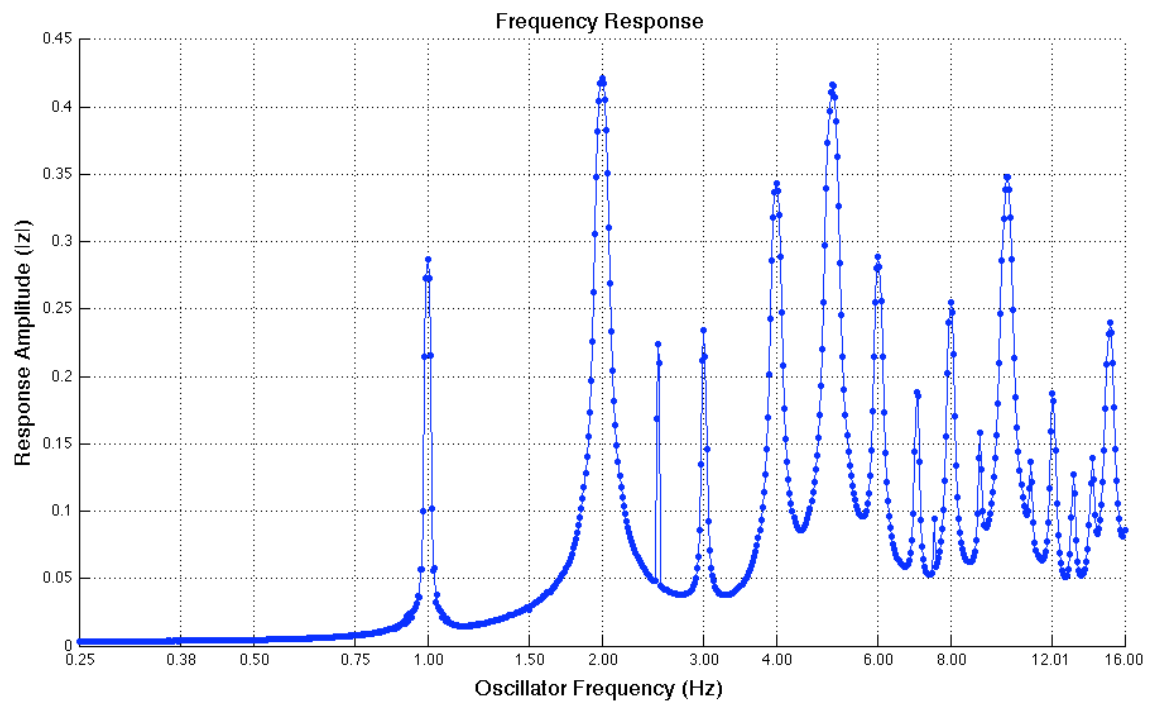


Figure 6.14: Frequency response of the model in the presence of a 2:5 polyrhythm repeating once every 1000 msec. The stimulus consists of a 2 Hz pulse train and a 5 Hz pulse train.



### 6.2.6 Simulation with a 3:5 polyrhythmic stimulus repeating once per second

The next simulation is concerned with the frequency response of the neural resonance model in the presence of a 3:5 polyrhythm repeating once per second. Therefore the input frequencies to the model are two pulse trains of 3 Hz and 5 Hz. In Figure 6.15 below we can observe that the frequency responses are structurally similar to the responses in previous simulations. So for example, we get responses that resemble the input frequencies of 3 Hz and 5 Hz, harmonics of those frequencies, e.g., 6 Hz, 9 Hz, and 12 Hz (harmonics of the 3 Hz input frequency) and 10 Hz and 15 Hz (harmonics of the 5 Hz input frequency), the first subharmonics of the input frequencies, i.e., 1.5 Hz and 2.5 Hz, a subharmonic response at 1Hz, which implies the frequency repetition of the pattern, and responses at 2 Hz and 7 Hz as a result of subtracting and summing the input frequencies respectively.

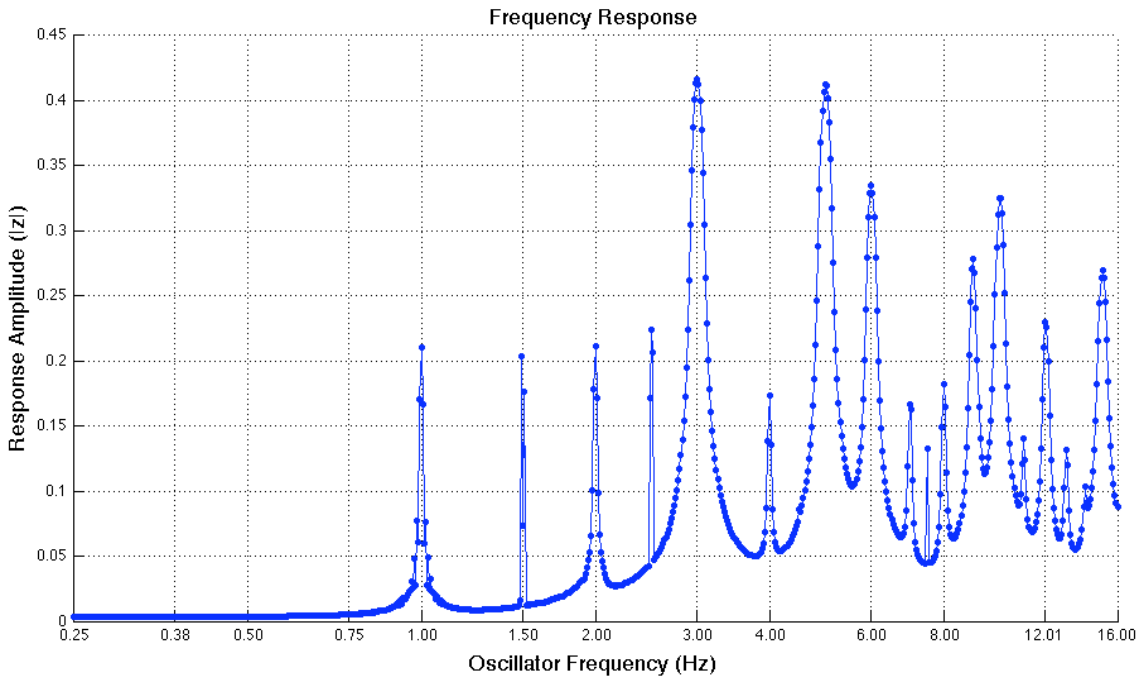


Figure 6.15: Frequency response of the model in the presence of a 3:5 polyrhythm repeating once every 1000 msec. The stimulus consists of a 3 Hz pulse train and a 5 Hz 5-pulse train.

The empirical findings reported by Handel and Oshinsky (1981) in regard to how people tap along with a 3:5 polyrhythm fall within the main categories reported by Handel and Oshinsky (in terms of preferences) of tapping along with either one of the pulse trains or tapping along with the co-occurrence of the pulse trains. In Figure 6.15 the aforementioned tapping behaviours are represented by frequency responses at 3 Hz, 5 Hz and 1 Hz respectively.

#### 6.2.7 *Simulation with a 4:5 polyrhythmic stimulus repeating once per second*

The next simulation concerns the last type of polyrhythm used in the Handel and Oshinsky study. This is a 4:5 polyrhythm repeating once per second so the input frequencies are 4 Hz and 5 Hz. The frequency response pattern can be explained in a similar manner to the analysis of the previous simulations. Therefore, regardless the type of polyrhythm presented to the neural resonance model the pattern of responses is structurally similar. This gives us the opportunity to generalise the type of expected responses in cases where the neural resonance model is stimulated by a polyrhythmic stimulus of two pulse trains. We develop the idea below and the frequency response pattern related to the 4:5 polyrhythm (Figure 6.16) is used to support the generalisation.

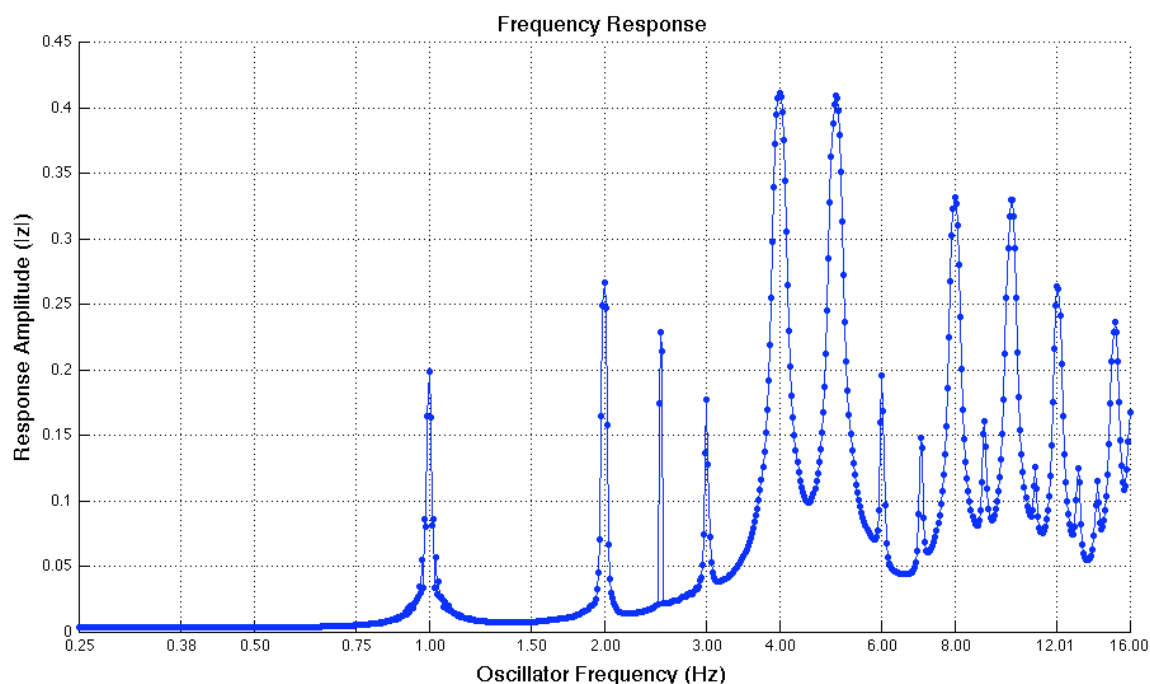


Figure 6.16: Frequency response of the model in the presence of a 4:5 polyrhythm repeating once every 1000 msec. The stimulus consists of a 4 Hz pulse train and a 5 Hz pulse train.

### 6.3 GENERALISING THE FREQUENCY RESPONSE PATTERN

Assuming that the input frequencies of a polyrhythmic stimulus are expressed as  $f_1$  and  $f_2$  respectively, it is expected that the neural resonance model would resonate at the following frequencies in the presence of any polyrhythmic stimulus consisting of two pulse trains (Table 6.2).

Table 6.2: A generic frequency response pattern of the neural resonance model in the presence of a two-pulse train polyrhythm.

a) Input frequencies $f_1$ and $f_2$
b) Harmonics of input frequencies, e.g., $2f_1$ , $3f_1$ , $4f_1$ , $2f_2$ , $3f_2$ , etc.
c) 1st subharmonics of the $f_1$ and $f_2$ , i.e., $f_1/2$ , $f_2/2$
d) Repetition frequency of the polyrhythmic pattern
e) Responses at $f_2-f_1$ and $f_2+f_1$

Figure 6.16 above contains responses that can be interpreted on the basis of Table 6.2. The above generalization, in particular, cases a), c), and d) resemble the general framework of tapping behaviours related to polyrhythm perception reported in the Handel and Oshinsky study. Furthermore, cases b), and e) in Table 6.2 suggest a possible extension of the above general framework, i.e., the ways people perceive beat or meter in polyrhythms. We have already discussed that the frequency responses falling within categories b) and e) can be attainable by people based on the upper and lower limits of sensorimotor synchronization tasks reported in the literature and also imply the ability of people to perceive metrical structure in a given rhythm (see 6.1.2).

The suggested extension as this derives from the general behaviour of the neural resonance model in the presence of a polyrhythmic stimulus can be tested by conducting experiments with people similar to those done by Handel and Oshinsky. In fact, the suggested extension is indicative of a very important process in scientific discovery, which is the use of a model to make predictions about human behaviour previously unobserved. We have valued this opportunity and for that reason we conducted experiments to examine the validity of the aforementioned predictions. The results of these experiments are discussed extensively in Chapter 8.

#### 6.4 SIMULATIONS WITH NON-ISOCHRONOUS POLYRHYTHMIC STIMULI

In modelling rhythm perception, the task of accommodating stimuli that deviate from perfect isochrony has been described as one of the most

troublesome to deal with (Large, 2008). In Chapter 4, we have discussed a case whereby the neural resonance model is stimulated with a freely improvised melody in order to test whether the beat of the melody can be identified. Indeed, the frequency response pattern gives rise to resonances, which correspond to the periodicity levels of quarter and half note. Given the focus of this thesis in polyrhythmic stimuli we wanted to stimulate the model with polyrhythmic stimuli performed by a single person in order to examine further the competence of the neural resonance model to deal with polyrhythmic stimuli of non-isochronous pulse trains. The frequency response pattern of the neural resonance in the presence of a polyrhythm with non-isochronous pulse trains can be compared with a pattern related to the same polyrhythm but with isochronous pulse trains.

An additional motivation for examining polyrhythms made of non-isochronous pulse trains is the relevance of such stimuli with real music and consequently the opportunity of testing the neural resonance model with stimuli that strive to resemble real stimuli as closely as possible. For example, much African music is fundamentally polyrhythmic and it is performed live on demand in order to mark a wide range of social occasions related to everyday life. Performing live means that a perfectly isochronous polyrhythmic stimulus like those we considered so far in this chapter is unlikely to be achieved by humans. In Chapter 2 we discussed that temporal fluctuations from perfect isochrony occur as a result of stylistic decisions but also because of people's natural tendency to deviate from perfect isochrony even if they are asked to tap periodically. Despite this deviation from perfect regularity in musical structure people can still perceive an underlying periodicity. A typical behavioural manifestation of perceiving regularity in African polyrhythmic music is dance. Dancing gestures can be triggered by spoken word, vocal music, and instrumental

music (Agawu, 1995). Dancing gestures in African music exhibit several types of synchronisation to the metrical structure of polyrhythmic music. For example, one such type of synchronisation is in-phase synchronisation with one particular rhythm. Therefore, the perception and the embodiment of rhythm as these are manifested through dancing are associated with the presence of polyrhythmic stimuli that exhibit temporal fluctuations - with regard to a perfect isochrony - in their metrical structure.

#### 6.4.1 *Preparation of polyrhythmic stimulus comprising non-isochronous pulse trains*

Two different polyrhythmic stimuli were used to stimulate the neural resonance model. We asked a performer to produce those two polyrhythms in the following way. In the first case, the polyrhythm is performed in the presence of a metronomic click, thus tempo variation is restricted but local deviations from perfect isochrony are expected - given that a musician performs the polyrhythm. In the second case, the polyrhythm is performed in the absence of a metronomic click (after four initial clicks), which allows for tempo variation over time as well as local deviations from perfect isochrony. For convenience of presentation, we asked the musician to perform a 4:3 polyrhythm at 60 bpm tempo, i.e., the polyrhythmic pattern is asked to be repeated several times once per second. We also asked the musician to tap the polyrhythm on a midi keyboard in order to record his performance. More specifically, the polyrhythm was performed bimanually, i.e., the right hand index finger playing the 4-pulse train and the left hand index finger playing the 3-pulse train.

Figure 6.17 is an excerpt of a time series representation of the performed polyrhythm, which also illustrates the form of the signal that is fed into

the neural resonance model. In this case, the polyrhythm is performed while the performer is listening to the metronome. The strongest response corresponds to the co-occurrence of the two pulse trains, i.e., the 1<sup>st</sup> pulse of both pulse trains. Then there are 5 events between the two successive peak responses. The 1<sup>st</sup>, 3<sup>rd</sup> and 5<sup>th</sup> of those are the 2<sup>nd</sup>, 3<sup>rd</sup>, and 4<sup>th</sup> pulses of the 4-pulse train. The 2<sup>nd</sup> and 4<sup>th</sup> events correspond to the 2<sup>nd</sup> and 3<sup>rd</sup> pulse of the 3-pulse train. The most obvious temporal fluctuation in Figure 6.17 is that the polyrhythm is drifting slightly to the right gradually over time, i.e., there is a deceleration, which indicates a delayed start of repeating the pattern with regard to the one-second period indicated by the inter-onset interval of the metronomic click. Additional fluctuation may exist between successive events like for example the relative position of the 4<sup>th</sup> event amid the 3<sup>rd</sup> and 5<sup>th</sup> events.

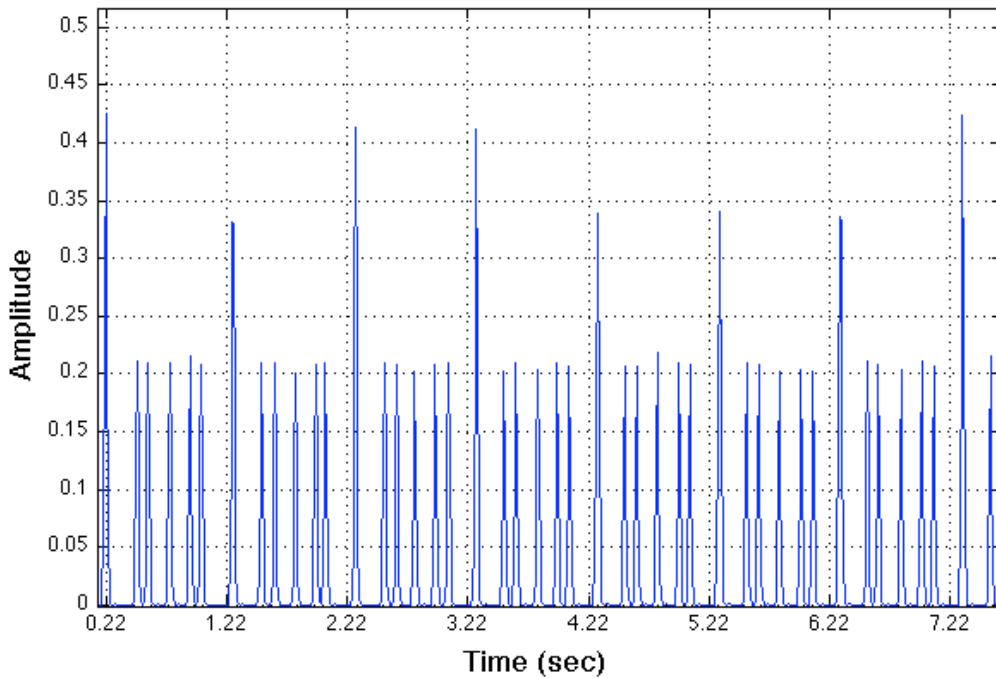


Figure 6.17: Time series representation of a 4:3 polyrhythm performed by a human at 60 bpm.

#### 6.4.2 *Stimulating the model*

Figure 6.18 below is the frequency response pattern of the neural resonance model in the presence of the above 4:3 polyrhythmic stimulus. We can see that despite the temporal fluctuation from perfect isochrony embedded in the pulse trains of the polyrhythmic stimulus the pattern is structurally similar to the one related to the stimulation of the model with a polyrhythm comprised of isochronous pulse trains. In other words, the frequency response pattern conforms to the generalised framework of the model's responses suggested previously (see Table 6.2). The only difference is a missing response corresponding to the 1<sup>st</sup> subharmonic of the 3-pulse train. This is likely due to the form of stimulus encoding of the human performance (MIDI). In Chapter 7 we will see that this particular response is significantly decreased also for the isochronous polyrhythm when the type of the encoding is a MIDI file.



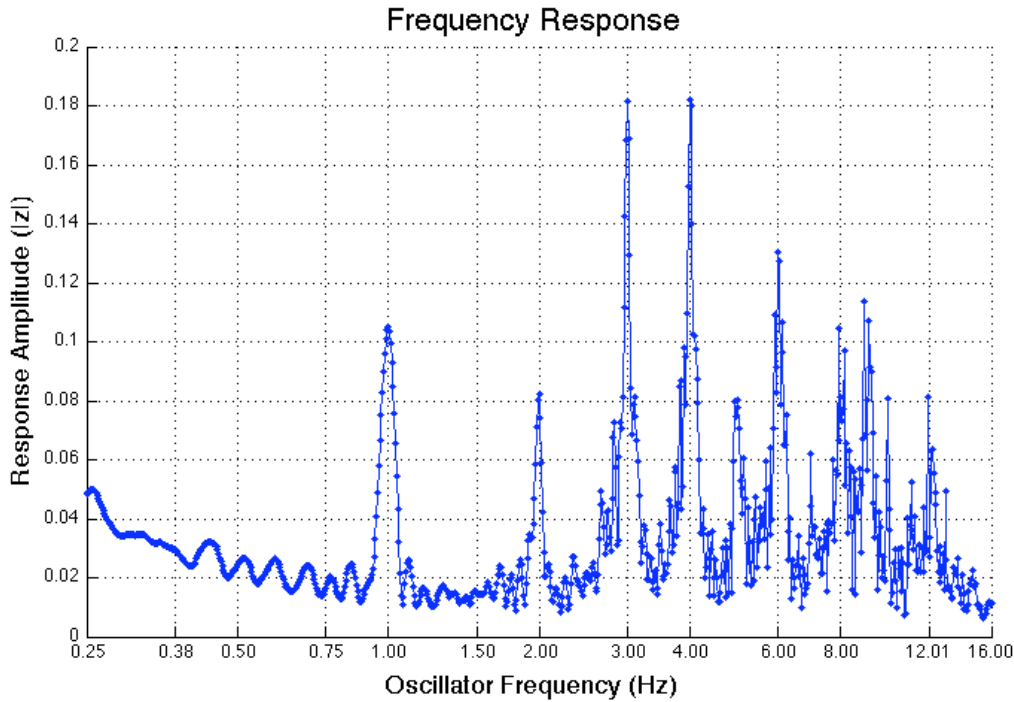


Figure 6.18: Frequency response of the model in the presence of a 4:3 polyrhythm performed by a human at 60 bpm.

The following simulation is the last of this chapter and is similar to the previous one, however the polyrhythm was performed without listening to the metronome, thus the polyrhythmic stimulus is susceptible to exhibit tempo variation as well as local fluctuations from perfect isochrony. In particular, in the current case four clicks at 60 bpm were provided in advance to the performance of the 4:3 polyrhythm and then the musician performed the polyrhythm for about 30 seconds without listening to the metronome. A time series representation of the first 9.5 seconds of the performed polyrhythmic signal is shown in Figure 6.19. This excerpt illustrates the temporal deviation occurring in the performance. In particular if we concentrate on the strongest amplitude response, which corresponds to the co-occurrence of the two pulse-trains, we can see that the second and third repetition of the polyrhythmic pattern is ahead of a hypothetical metronomic click occurring every second. If the first

hypothetical click is taken to be at 0.5 seconds in the x-axis, then all numerical points in the x-axis demarcate a hypothetical metronomic click. The performance starts approximately at 500 msec, a point in time that marks the first co-occurrence of the two pulse trains. It is therefore expected that the second repetition of the polyrhythmic pattern will start at 1.5 sec, insofar as the performer was asked to perform the polyrhythm at 60 bpm, i.e., to repeat the pattern once per second. However, we can see in the graph that the second repetition of the polyrhythmic pattern starts slightly ahead of time, i.e., the first repetition was a bit too fast. During the second repetition it seems like the performer may have felt that the first repetition was too fast so he is holding the former longer than 1 second in order to return to a one repetition per second mode. Despite this correction, the third repetition starts slightly delayed with respect to a hypothetical click at 2.5 seconds and also it lasts a bit longer than 1 second - maybe indicating a continuation of the previous correction strategy. Over the next two repetitions (4<sup>th</sup> and 5<sup>th</sup>) the polyrhythmic pattern is in phase with a hypothetical click and therefore the one repetition per second is successful. From repetition 6 onwards the performance starts drifting slightly to the right with respect to the hypothetical clicks, which means that the repetition of the polyrhythmic pattern is slower than expected. This performance is a typical example of temporal deviation from perfect isochrony and it shows the kinds of temporal fluctuation related to human performance discussed in Chapter 2 (see Subsection 2.1.3). In this last simulation the neural resonance model has been stimulated with the first 15 seconds of the above performance. Figure 6.20 corresponds to the frequency response pattern of the model averaged over the last 20% of the simulation time, i.e., the last 3 seconds. We can see that the pattern is similar to the two previous cases where a 4:3 polyrhythmic stimulus with repetition rate of 1 second

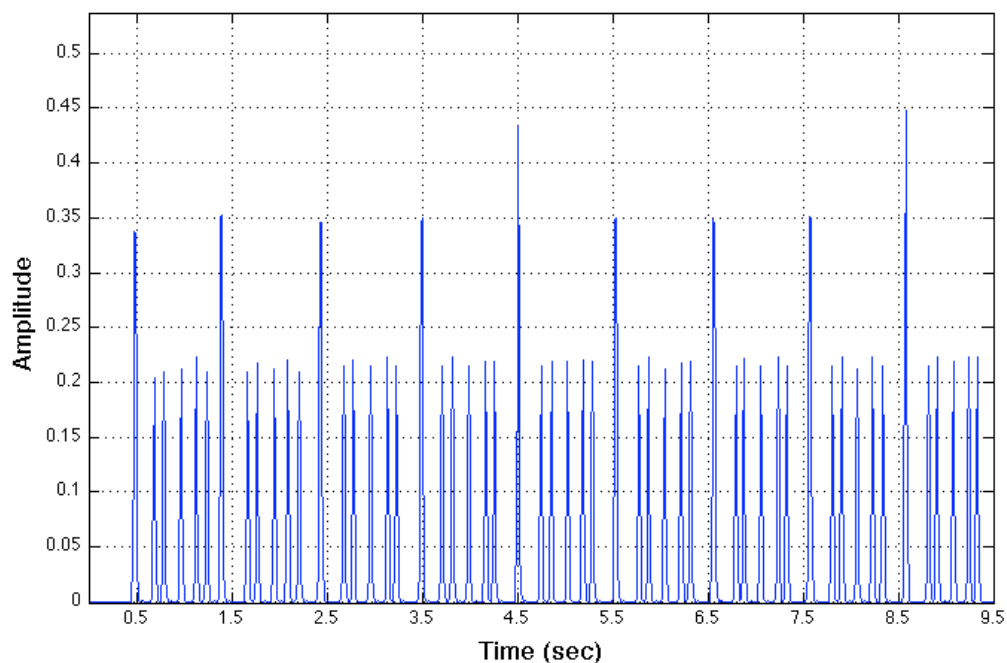


Figure 6.19: Time series representation of a 4:3 polyrhythm performed by a human at 60 bpm with no metronomic guidance.

comprised of isochronous pulse trains was presented to the model, and also, it conforms to the generalized framework suggested in Table 6.2.

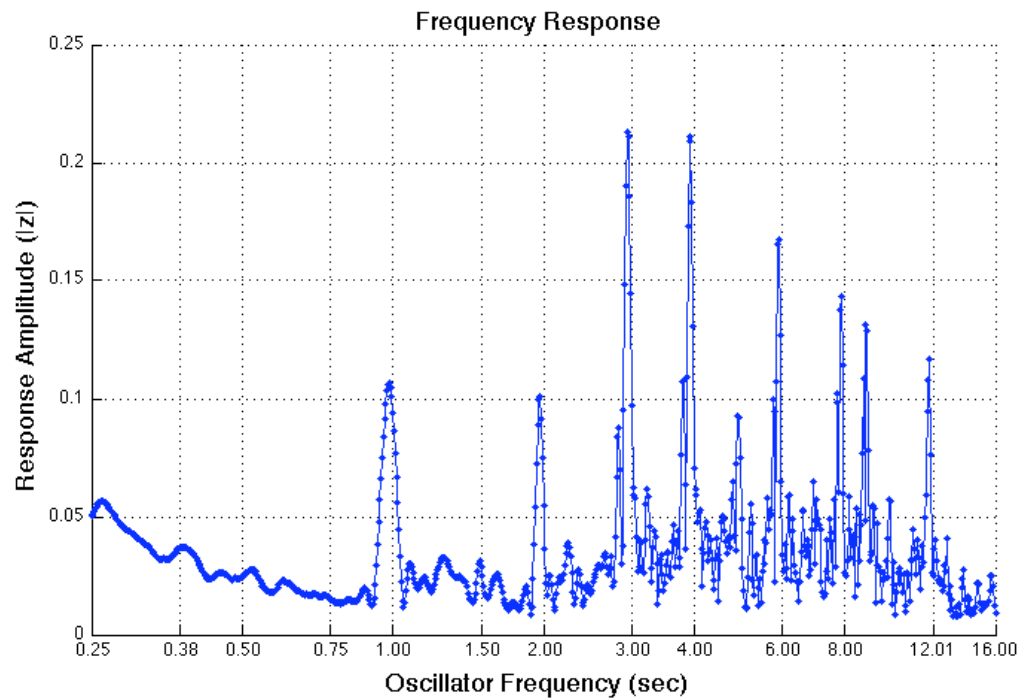


Figure 6.20: Frequency response of the model in the presence of a 4:3 polyrhythm performed by a human at 60 bpm with no metronomic guidance.

## 6.5 SUMMARY

Overall, based on the output of all the simulations presented in this chapter we can see that the neural resonance model produces responses that can be matched to tapping behaviours observed by Handel and Oshinsky (1981) for a range of polyrhythms and repetition rates. Also, the model can accommodate temporally variant polyrhythmic stimuli that are closer to the kind of rhythmical stimuli occurring in performed music. A general framework to capture the frequency response pattern of the neural resonance in the presence of a two pulse-train polyrhythm has been suggested. Inspection of the latter in combination with empirical studies about the upper and lower limits in sensorimotor synchronisation tasks suggest that there is an opportunity for conducting empirical experiments in order to potentially observe attainable tapping frequencies, which have

not been reported in the literature previously. Finally, the hypothesis that the magnitude of the amplitude responses in the model may be indicative of how relative tapping preferences are forming through transposition of the polyrhythm in time did not verify given that the frequency response pattern remains invariant for different tempi.

## MODEL ANALYSIS

---

In general, the process of validating mathematical models of biological systems involves the comparison between modelled data and empirical data. Furthermore, the extent to which a model manages to produce data similar to those empirically observed serves as a measure of the degree of validity of a model, subject to certain provisos. For example, the process of obtaining simulated and empirical data is in general accompanied by a degree of uncertainty regarding the adequacy of the collected data, on both empirical and simulated sides, to support inferences about the validity of a model. Some of the main sources of uncertainty when gathering data from a model to be tested, are related to the fact that any simulated dataset will be based on modelling assumptions about the natural system to be modelled. These assumptions may be incomplete or wrong (O'Neill and Gardner, 1979). In the case of a parameterised model, further issues may arise, since any simulated dataset may depend on specific numerical values of the parameters.

In Chapter 5 (see Subsection 5.2.6) we have briefly mentioned that the set of nominal values of the parameters used for testing of the neural resonance model could be seen as a potential source of uncertainty with regard to the obtained modelled data. Another such source of uncertainty relates to the issue that the neural resonance model is a canonical representation of the more biological Wilson-Cowan model. In order to address these and other issues, we will now discuss model validation techniques designed to

limit uncertainty. For example, such techniques include testing a model for consistency in producing certain results for a range of numerical settings, and comparing the performance of a model with the performance of other similarly purposed models. In general, similar-purpose models typically simulate the same biological system but they may make slightly different biological hypotheses and/or use different mathematical formulae. In the present research, both of the above two techniques for limiting uncertainty about the assumptions behind the neural resonance model are used. Firstly, a technique called parameter sensitivity analysis is employed, under which the numerical values of the parameters of the neural resonance model (detailed below) are systematically manipulated. Secondly, a technique called cross-model comparison is employed to limit uncertainty. In this case, the behaviour of the neural resonance model is compared with the behaviour of the Wilson-Cowan model. The Wilson-Cowan model is a well-respected and highly biologically plausible model for capturing the dynamics of neural populations (Wilson and Cowan, 1972). Thus, by comparing its behaviour with the behaviour of the neural resonance model we can limit uncertainty regarding the biological assumptions made by the latter.

### 7.1 PARAMETER SENSITIVITY ANALYSIS

In the next few sections, sensitivity analysis is conducted on a range of parameters of the neural resonance model. The neural resonance model contains many different parameters. In particular, parameters appear in several places such as in the *state variable* of the neural resonance model, i.e., the equation that describes a neural oscillator, in the *driving variable*, i.e.

the *external stimulus*, and in the *initial conditions*, i.e., *initial amplitude* and *phase* of the oscillators.

The state variable parameters (see Chapter 4 - Equation 7) account for the dynamics of a neural oscillator over time in the presence of a rhythmic stimulus. The external stimulus is an independent variable (driving variable) that influences the dynamics of the state variable. The only parameter associated with the external stimulus is the amplitude of the latter, but analysis can also be performed for different types of stimulus encoding, e.g., MIDI vs. audio files (see Chapter 5 - Subsection 5.2.1). The initial conditions are related to the amplitude and phase value of each oscillator in the bank just before the stimulation starts. The next three subsections describe in detail the systematic manipulation of the numerical values for all types of parameters briefly introduced above, starting with the initial conditions of the oscillators, then the external stimulus, and finally, the parameters of the state variable of the neural resonance model.

#### 7.1.1 *Initial conditions of each oscillator: amplitude and phase*

Initial conditions refer to the initial numerical values given to the observable outputs of the state variable of a model at the starting point of a simulation. In the case of the neural resonance model these outputs are the amplitude and the phase of the oscillators comprising the bank. For example, the output of the simulations presented in Chapter 6 is the average value of the amplitude of each oscillator over the last 20% of the stimulation with an external polyrhythm. From that perspective, initial numerical values should be given to both the amplitude and the phase of each oscillator in the bank at time zero. As a result, we need to identify those initial numerical values, and to clarify why a certain numerical value



is chosen instead of any other. To address this point, at least in the context of this thesis, it is useful to remember that we are mostly concerned with examining the amplitude responses of the neural resonance model in the presence of some stimulus. That is we want to examine which oscillators in the bank will resonate in the presence of some polyrhythmic stimulus. The resonance phenomenon can briefly be described as the increase in amplitude of an oscillation of a physical system due to the exposure to a periodic external force of which the driving frequency is equal or very close to the natural frequency of the physical system (van Noorden and Moelants, 1999). In effect, the focus is on the increase of the amplitude value of each oscillator resulting from resonance activity, i.e., only one of the two main metrics of the state variable of the neural resonance model (the other being the phase), thus the discussion below is exclusively focused on the initial condition of the amplitude. With regard to the phase of the oscillators, all oscillators are in-phase in the beginning of a simulation, i.e., relative phase between any pair of oscillators in the bank is zero.

In the presence of a polyrhythmic stimulus some of the oscillators in the bank are expected to resonate (exhibit an increase in their amplitude) regardless of their initial amplitude. One reasonable numerical setting for the initial amplitude of each oscillator in the bank is a value close to zero, a numerical value that is used to indicate baseline activity, i.e., neural activity in the absence of external stimuli. In fact, the value of the initial amplitude of the oscillators in the simulations presented in Chapter 6 is indeed close to zero  $10^{-10}$ . However, the response of the neural resonance model can be examined for a numerical range regarding the initial amplitude of the oscillators such as between 0 (inclusive) and 1 (exclusive). For any value within this range we would expect to observe resonance activity similar to this presented in Chapter 6. Indeed, the following three figures illustrate

that the resonance effect is independent of the initial amplitude of the oscillators, and that the initial amplitude of the oscillators only corresponds to a transient state of the system. Figure 7.1 and Figure 7.2 illustrate the effect of resonance for initial amplitudes 0.7 and  $10^{-10}$  respectively in a 12.5 second stimulation period using a polyrhythmic stimulus consisting of two squarewaves 3Hz and 4Hz. Figure 7.3 illustrates the combined results of the average activity of the bank of oscillators over the last 20% of the stimulation.

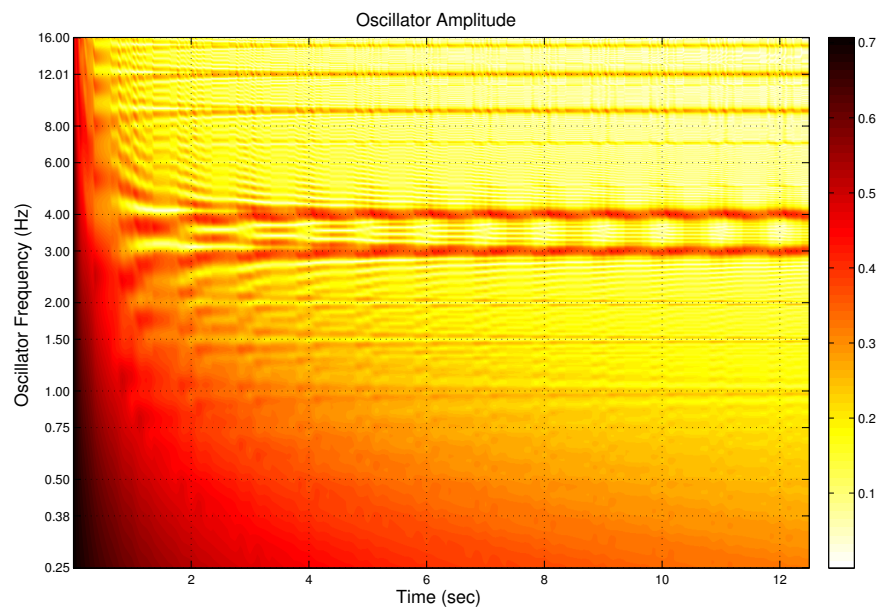


Figure 7.1: Amplitude response of the neural resonance model over a period of 12.5 seconds where the initial amplitude for all oscillators is 0.7. Gradually, the oscillators drop the levels of the initial amplitude except from those oscillators that resonate in the presence of the polyrhythmic stimulus. Some of the resonating oscillators, which are easy to spot, have the following frequencies: 12 Hz, 9 Hz, 4 Hz, 3 Hz, 2 Hz, 1.5 Hz, and 1 Hz.

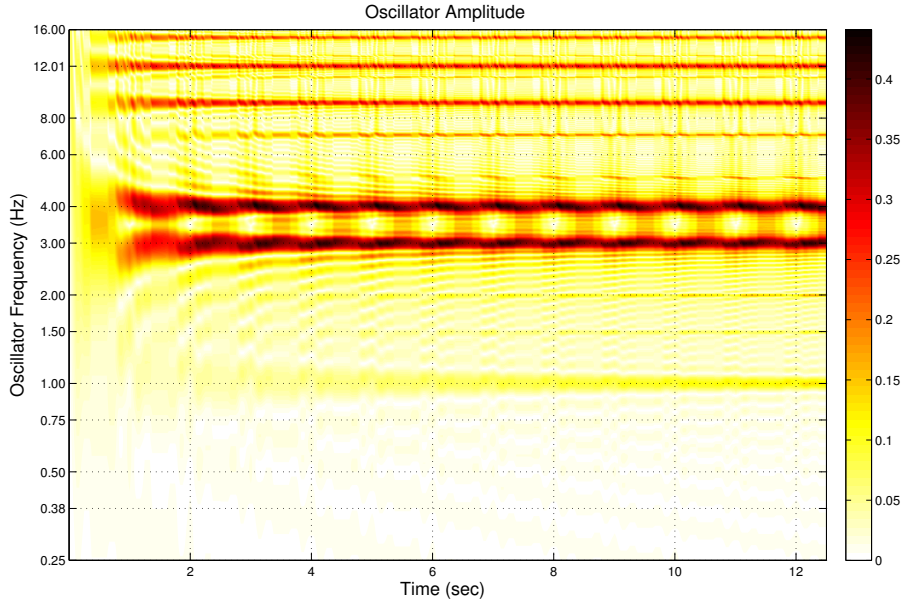


Figure 7.2: Amplitude response of the neural resonance model over a period of 12.5 seconds where the initial amplitude for all oscillators is close to 0. Over time, the oscillators that resonate in the presence of the polyrhythmic stimulus exhibit higher levels of amplitude. Some of the resonating oscillators, which are easy to spot, have the following frequencies: 12 Hz, 9 Hz, 7 Hz, 4 Hz, 3 Hz, and 1 Hz.

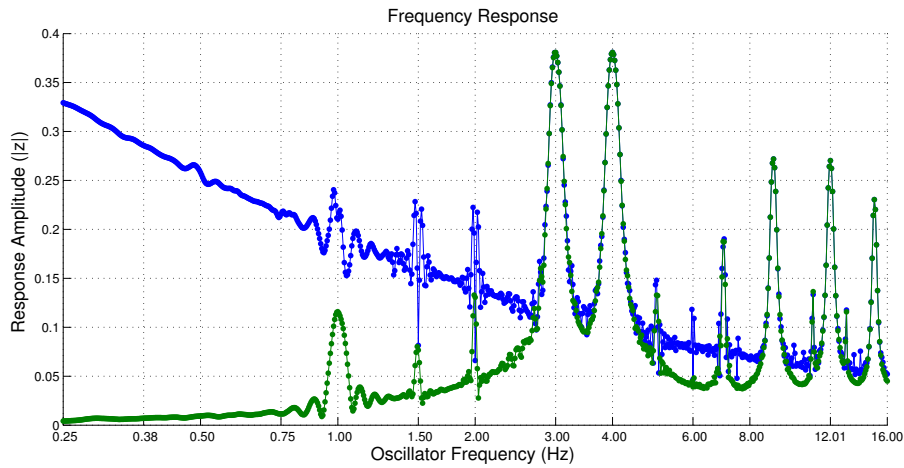


Figure 7.3: Superimposed frequency response patterns for different initial amplitude of the oscillators 0.5 (blue) and near 0 (green). The resonance pattern is similar in both cases, i.e., similar oscillators resonate in the presence of the polyrhythmic stimulus regardless the initial amplitude. The high amplitude levels of the low frequency oscillators in the blue pattern correspond to a transient stage.

### 7.1.2 *Driving variable parameters: external polyrhythmic stimulus*

In general, driving variables are independent variables, which are used to influence a state variable. They are dynamic variables acting as an input to a state variable. Driving variables have no input. Systematic numerical manipulation of the parameters of a driving variable allows us to examine their effect in the general behaviour of a model, and the consistency of the latter in producing qualitatively similar results. In the simulations presented in Chapter 6 this type of driving variable is the external polyrhythmic stimuli used to stimulate the neural resonance model. The parameter we are concerned with here is the amplitude of the stimulus. However, the different ways of encoding the external stimulus, e.g., MIDI and audio files (see Chapter 5 - Subsection 5.2.1) can be subjects for sensitivity analysis as well. The results of applying parameter sensitivity analysis with regard to the amplitude of the stimulus and its encoding type are presented below.

In Chapter 6 the polyrhythmic stimulus was encoded using functions. Figure 7.4 below combines frequency response patterns of the neural resonance model for all three potential ways of stimulus encoding i.e., functions, audio files, and MIDI files. The fact that the frequency response patterns of the neural resonance model are structurally the same for all the above ways of stimulus encoding indicates that the model behaves consistently regardless the chosen way of stimulus encoding. The blue frequency response pattern corresponds to using the bump function for encoding the stimulus, the green to using audio files, and the red to using MIDI files for encoding the stimulus. By comparing the frequency response patterns we can see that the peak responses are similar for all the patterns regardless the encoding technique. However, we should note the absence

of even harmonics in the audio encoding which uses square waves to represent the polyrhythm. In Chapter 6 we mentioned that this is normal since the decomposition of a squarewave into an infinite set of simple oscillating functions only contains components of odd integer harmonic frequencies.

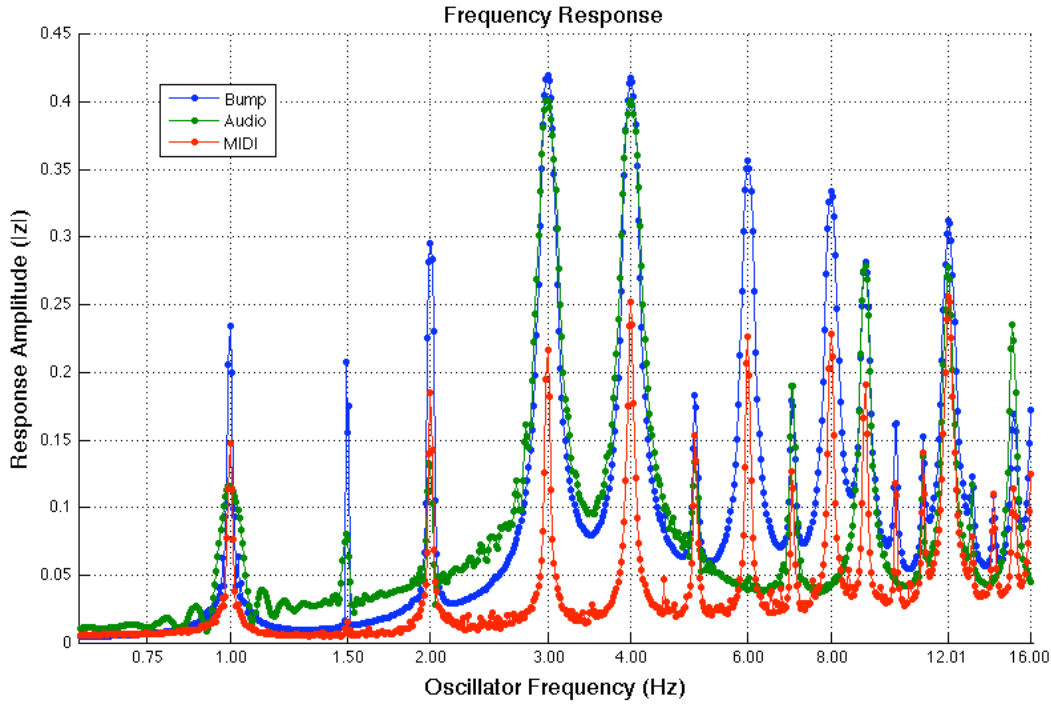


Figure 7.4: Superimposed frequency response patterns of the neural resonance model for the three different ways of encoding a polyrhythmic stimulus, i.e., audio files, MIDI files and bump function.

We now proceed to examine what happens to the frequency response pattern of the bank by considering a systematic manipulation of the amplitude of the polyrhythmic stimulus. The stimulus is encoded using the bump function because it takes less time to alter the amplitude value compared to the other ways of encoding. The amplitude value of the stimulus is manipulated around a central value (in this example 0.4) within a certain range (0.2 - 0.6). These values can be defined arbitrarily, taking into consideration certain limitations on the stimulus amplitude imposed

by the mathematical derivation of the model (Large et al., 2010). Figure 7.5 below illustrates the frequency response patterns for three numerical values of the stimulus's amplitude. The red response corresponds to the largest amplitude (0.6), the blue to the medium one (0.4), and the green to the smallest amplitude (0.2). In general, we can see that the magnitude of the response of the oscillators follows the changes in the amplitude of the stimulus, but nonetheless the resonance pattern is similar to all three cases.

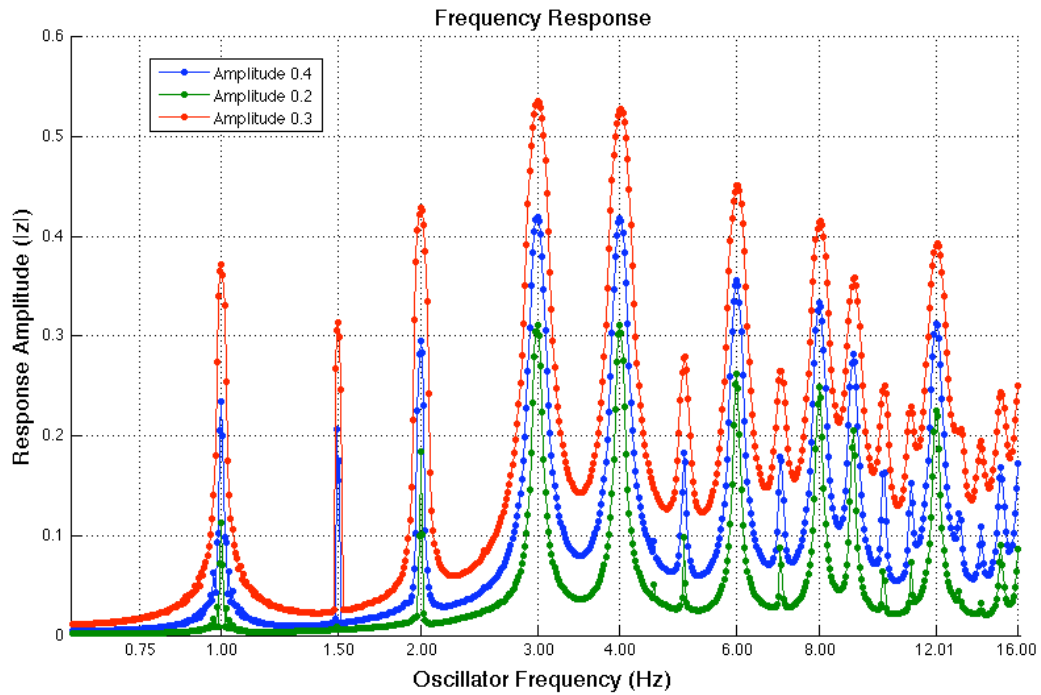


Figure 7.5: Superimposed frequency response patterns for different amplitudes of a 4:3 polyrhythmic pattern with 1 Hz repetition frequency.

In general, we can see that the frequency response patterns are structurally the same regardless the differences in the amplitude of the stimulus and its encoding type. Thus, the behaviour of the neural resonance model is consistent in producing similar results to those presented in Chapter 6 regardless any small variations in the value of the amplitude of the stimulus and its encoding type.

### 7.1.3 *State variable parameters*

In this section we apply sensitivity analysis to the parameters of the state variable of the neural resonance model, which is represented by Equation 7 (Chapter 4). The equation will be repeated below to facilitate the discussion and illustrate the parameters. Similarly to the previous sections, the generally desirable expectation is that a slight change in the parameter values will not create significant changes in the frequency response pattern of the neural resonance model. The idea that a systematic, but limited in numerical range, manipulation of the values of the parameters of a model will not cause a significant change in the responses of the model, is expected due to an intuitive belief that most real systems will not respond violently to small changes in the values of the operating parameters or variables (Haefner, 2005). The section begins by disclosing the nominal numerical values given in the parameters in the simulations of Chapter 6, and continues by applying sensitivity analysis to some of the parameters comprising the state variable of the neural resonance model.

#### 7.1.3.1 *Nominal values of state variable parameters*

In Chapter 4 the mathematical formulation of the neural resonance model was presented and the functionality of each parameter has been described. However, at that point we did not disclose the numerical values of the parameters during a simulation. Also, in Chapter 5 we briefly mentioned that the nominal values of the parameters of Equation 7 that represents a neural oscillator are defined based on certain criteria such as the fact that we are interested in the dynamics of an oscillator near an Andronov-Hopf bifurcation point. There we pointed forward to this chapter as a more appropriate place to include a detailed discussion on the nominal values

of the parameters, because parameter sensitivity analysis deals exclusively with manipulation of numerical values of parameters.

The following equation is Equation 7 introduced in Chapter 4. Parameters  $b$ ,  $d$ , and  $k$ , are complex variables (disclosed below), which reveal the full range of parameters associated with the state variable ( $z$ ) of the neural resonance model.

$$\dot{z}_i = b_i z_i + d_i z_i |z_i|^2 + \frac{k_i \varepsilon z_i |z_i|^4}{1 - \varepsilon |z_i|^2} + \frac{x_i(t)}{1 - \sqrt{\varepsilon} x_i(t)} \cdot \frac{1}{1 - \sqrt{\varepsilon} x_i(t)} \quad (8)$$

where,

$$b = \alpha + i\omega$$

$$d = \beta_1 + i\delta_1$$

$$k = \beta_2 + i\delta_2$$

Thus, the range of parameters associated with the neural resonance model is  $(\alpha, \beta_1, \beta_2, \delta_1, \delta_2, \omega)$  and  $(\varepsilon)$ . Parameters  $(\alpha, \beta_1, \beta_2, \delta_1, \delta_2)$ , and  $(\omega)$  relate to the oscillatory nature of the system. Parameter  $(\alpha)$  corresponds to a bifurcation parameter, which is in effect a parameter denoting the systems transition from one dynamic state to another. In this thesis, the bifurcation parameter takes a nominal value equal to zero, which implies that the system lies at a critical point called Andronov-Hopf bifurcation point. As previously mentioned the reason for taking parameter  $(\alpha)$  equal to 0 is related to the dynamics of the original system of equations describing the interaction of excitatory and inhibitory neural populations. When a dynamical system like the above is near one of its bifurcation points (e.g., Andronov-Hopf) the interaction of the excitatory and inhibitory populations may result either in a steady oscillation or a damped oscillation, i.e., the neural oscillator is in a transient mode. For



positive values, the oscillatory system moves on to spontaneous oscillation while for negative values it moves on to a damped oscillation. Sensitivity analysis for parameter ( $\alpha$ ) can therefore be applied by considering a systematic numerical manipulation within a range, which has its centre around the nominal value zero.

Parameter ( $\beta_1$ ) and ( $\beta_2$ ) are damping parameters that prevent the amplitude of the oscillation from becoming infinitely large when ( $\alpha$ ) is positive. In other words, they function as a saturation point (limit) around which the amplitude of the oscillation is bound to stabilise. In the simulations of Chapter 6 some of the oscillators exhibit an increase in their amplitude due to the resonance effect, therefore parameters ( $\beta_1$ ) and ( $\beta_2$ ) control the resonance amplitude. The nominal values of parameters ( $\beta_1$ ) and ( $\beta_2$ ) should be negative to express the above functionality. In the simulations of Chapter 6, only parameter ( $\beta_1$ ) had a negative nominal value of -1, while ( $\beta_2$ ) was set to 0 for simplification. The numerical range employed for conducting sensitivity analysis with respect to ( $\beta_1$ ) is discussed in the following section.

Parameters ( $\delta_1$ ) and ( $\delta_2$ ) are frequency detuning parameters that define how the peaks in the frequency response pattern begin to detune as the strength of the stimulus increases. Frequency detuning was illustrated in Figure 4.6 (Chapter 4) for positive values of ( $\delta_1$ ) and ( $\delta_2$ ). Large (2008) suggested that frequency detuning - for negative values - can address some behavioural observations, such as the tendency of people to tap ahead in time when asked to tap along with a metronomic click. This tendency of tapping ahead of time is illustrated in Figure 7.11 where for example the maximum resonance for a pulse train of 3 Hz rhythmic frequency corresponds to an oscillator with natural frequency slightly larger than 3 Hz. In the simulations presented in Chapter 6 both ( $\delta_1$ ) and ( $\delta_2$ ) parameters

were set to zero for simplicity. We have already discussed how the sign of the value of parameters ( $\delta_1$ ) and ( $\delta_2$ ) influences the direction of detuning. For that reason we can infer that a parameter sensitivity analysis whereby the values of parameters ( $\delta_1$ ) and ( $\delta_2$ ) take a small increase or a small decrease around their nominal value zero would detune the frequency response pattern either to the left or to the right respectively.

Parameter ( $\omega$ ) is the angular frequency of the oscillator. It refers to the product of the natural frequency of each one of the oscillators in the bank multiplied by  $2\pi$ . The frequency of each oscillator is fixed, thus it makes no sense to apply sensitivity analysis for ( $\omega$ ) either.

Lastly, parameter ( $\varepsilon$ ) controls the type of coupling between oscillators and the external stimulus. When the value of ( $\varepsilon$ ) is 0 the coupling is linear, while for positive values it becomes nonlinear. Linear coupling implies that only the oscillators with natural frequencies close to those of the input frequencies of the stimulus and their harmonics will resonate. However, if ( $\varepsilon$ ) is set to 1 then there is a nonlinear coupling between the bank of oscillators and the stimulus, and the full range of resonances such as subharmonics and higher order combinations of the input frequencies emerges. In the simulations of Chapter 6 the nominal value of parameter ( $\varepsilon$ ) was 1, so parameter sensitivity analysis will be considered for a range of values around 1 but strictly non-zero in order to retain nonlinearity.

### 7.1.3.2 *Setting numerical range for state variable parameters*

Sensitivity analysis examines whether the systematic numerical manipulation of the parameters of the state variable described above has an effect in changing the frequency response pattern of the neural resonance model. In the previous section we revisited the parameters related to the state variable of the model, and we disclosed the nominal

numerical values that were used in the simulations of Chapter 6. Also, we discussed that sensitivity analysis will be applied only to some of those parameters, and in particular, to parameters  $(\alpha)$  with nominal value 0,  $(\beta_1)$  with nominal value  $-1$ , and  $(\varepsilon)$  with nominal value 1.

The sensitivity analysis entails running multiple simulations whereby numerical values are changed systematically for the three parameters mentioned above. The coefficients of similarity  $r^2$  of the frequency response patterns are then calculated. For each parameter we need to define a numerical range around the nominal value previously described. Additionally, we have to determine how many numerical values within each numerical range for each parameter are chosen for the analysis. For example, by choosing four values for each of the three parameters the total number of simulations to run will be  $4^3 = 64$ . Within this total number of combinations, some simulations will only have the value of a single parameter changed with respect to the nominal values, others will have the value of two parameters changed, and some others will have the values changed of all three parameters. The first category is described as to *single* parameter sensitivity analysis, whereas the last two are described as *multiple* parameter sensitivity analysis. In some cases, where the number of simulations is large the analysis can be confined by choosing multiple parameter analysis exclusively. Another way to bring down the total number of simulations to be run is by taking fewer values for each parameter within its range. The exact number of runs is discussed in the following sections.

We have already introduced the parameters of the state variable. The nominal values of these parameters were discussed along with some implication about the range of values for each parameter that can be potentially used in the parameter sensitivity analysis. In general, the

range of values for each parameter corresponds to the degree of alteration around the nominal value. There are two strategies for determining the amount by which parameters can be altered, namely the *uniform* and the *variable* approach (Haefner, 2005). In the *uniform* approach, all parameters are altered by the same percentage with respect to their nominal value. In order to apply the *variable* approach we need to know the variance of the nominal value, if any, for specific parameters. However, since we do not have information about the variance of nominal values for certain parameters regarding the neural resonance model, we can employ the uniform approach. Often the percentage of alteration in the uniform approach is  $\pm 10\%$  around the nominal value, but values ranging from 2% and 20% can also be used (Haefner, 2005). Table 7.1 below shows the numerical values of the parameters ( $\alpha$ ), ( $\beta$ ), and ( $\varepsilon$ ) after adopting a  $\pm 10\%$  uniform variation around the nominal value.

Table 7.1: Range of numerical values for parameters ( $\alpha$ ), ( $\beta$ ) and ( $\varepsilon$ ). Values are calculated based on a  $\pm 10\%$  uniform variation around nominal values.

	Parameter ( $\alpha$ )	Parameter ( $\beta$ )	Parameter ( $\varepsilon$ )
Nominal Value	0	-1	1
+10%	0.1	-0.9	1.1
-10%	-0.1	-1.1	0.9

The developers of the neural resonance model (Large et al., 2010) have made suggestions about the range of values for parameter ( $\alpha$ ). In particular, they suggest that parameter ( $\alpha$ ) can take values between 0 (inclusive) and a small positive such as 0.1. The reason for suggesting this range is to focus on spontaneous oscillation mode (for small positive), which can account for simulating the endogenous periodicity associated with rhythm perception in people. In other words, in the spontaneous oscillation mode,

once an oscillation is established it can be retained even in the absence of an external stimulus, which according to Large (2010) is a potential mechanism that explains how people are able to retain a perceived rhythm in periods of pauses of events or even in highly syncopated rhythms. Large's range differs from the general guidelines mentioned above, in particular where the lower limit of parameter ( $\alpha$ ) with nominal value 0 is  $-0.1$  ( $-10\%$  of the nominal value). However, there is no restriction in considering negative values for ( $\alpha$ ), because the polyrhythmic stimuli used in simulations of Chapter 6 are not syncopated and have no long pauses, thus requirement for the spontaneous oscillation mode exclusively is not imperative. In other words, considering the presence of a polyrhythmic stimulus as this has been described so far, in combination to the fact that even a damped oscillator ( $\alpha < 0$ ) has the ability to resonate, we can expect the frequency response of the model to be similar to the one where critical<sup>1</sup> ( $\alpha = 0$ ) or spontaneous oscillation ( $\alpha > 0$ ) is considered. Figure 7.6 below illustrates the responses of the neural resonance model at the critical Andronov-Hopf point ( $\alpha = 0$ ) and also, when the oscillator is damped ( $\alpha = -0.1$ ). Despite the fact that the two graphs are highly similar with  $r^2 = 0.9463$ , the responses at 1.5 Hz, 2 Hz, and 5 Hz are significantly lower in the damped mode. In a way we could say that the system in its damped oscillation mode suppresses its ability to resonate in the presence of a stimulus. This is reasonable as the magnitude of the damping has an inversely proportional effect to the magnitude of the resonance. After conducting a further run in which we set ( $\alpha = -1$ ), which is not depicted here, the resonances were clearly suppressed for all the oscillators.

---

<sup>1</sup> A critical point is a non-equilibrium transient state of a dynamical system that lies between two different stable dynamic modes.

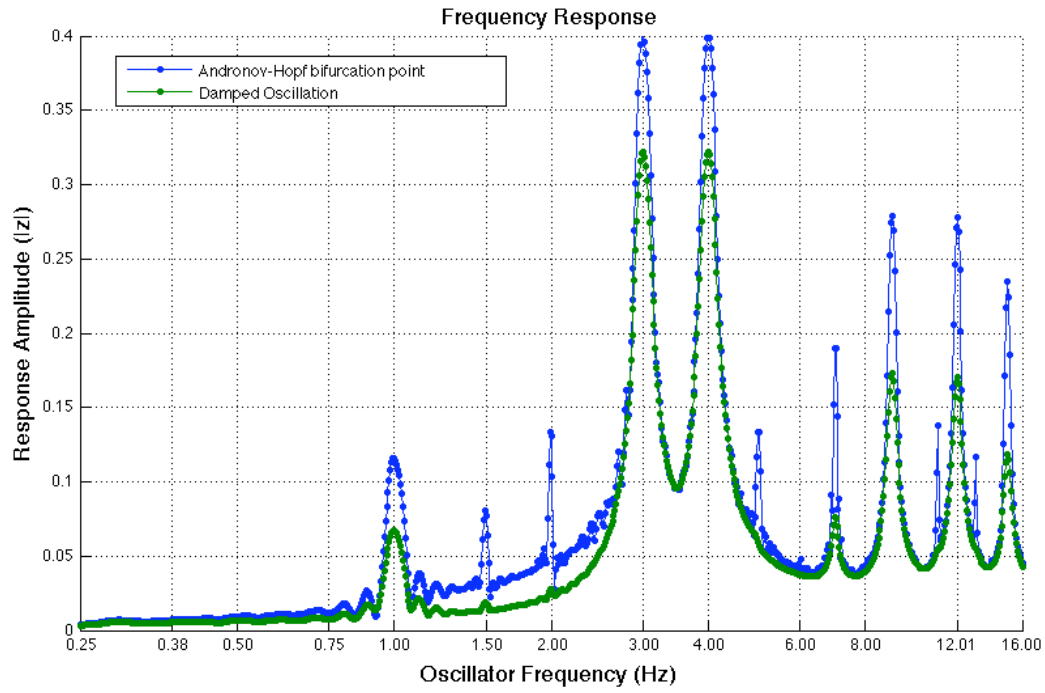


Figure 7.6: Frequency response patterns for ( $\alpha = 0$ ) (blue), and ( $\alpha = -0.1$ ) (green). The patterns are very similar with  $r^2 = 0.9463$

### 7.1.3.3 Sensitivity analysis for parameters ( $\alpha$ ) and ( $\epsilon$ )

In Figure 7.6 above we saw that even for a small value of the parameter ( $\alpha = -0.1$ ) the frequency response pattern starts dropping its resonances. Additionally, a large negative value of parameter ( $\alpha$ ) would eventually turn the system into a rigid non-responsive system in the presence of a stimulus. For that reason, we can take advantage of it and limit the total number of runs we have to perform in parameter sensitivity analysis by disregarding the cases whereby ( $\alpha$ ) takes a negative value. Table 7.2 below includes all possible runs for a ( $\pm 10\%$  uniform variation around the nominal value of parameters ( $\alpha$ ) and ( $\epsilon$ ) (negative ( $\alpha$ ) included).

Table 7.2: Combinations of different runs for parameters ( $\alpha$ ) and ( $\varepsilon$ ) considering both single and multiple sensitivity analysis.

Runs	Parameter ( $\alpha$ )	Parameter ( $\varepsilon$ )	Color
1	Nominal (0)	Nominal (1)	Blue
2	Nominal (0)	+10%(1.1)	Green
3	Nominal (0)	-10%(0.9)	Red
4	+10%(0.1)	Nominal (1)	Light Blue
5	+10%(0.1)	+10%(1.1)	Purple
6	+10%(0.1)	-10%(0.9)	Golden
7	-10%(-0.1)	Nominal (1)	n/a
8	-10%(-0.1)	-10%(0.9)	n/a
9	-10%(-0.1)	+10%(1.1)	n/a

Run 1 corresponds to the original simulations (Chapter 6) with nominal values ( $\alpha = 0$ ,  $\beta_1 = -1$ , and  $\varepsilon = 1$ ). Run 2, run 3, and run 4 imply single parameter analysis because only parameter ( $\varepsilon$ ) is altered in the first two cases, and only parameter ( $\alpha$ ) in the third case. Run 5 and run 6 imply multiple parameter analysis because two parameters are altered at the same time with respect to the nominal values. Figure 7.7 illustrates the frequency response patterns of the first six runs described in Table 7.2 in the presence of a 4:3 polyrhythmic stimulus with 1 Hz repetition frequency.

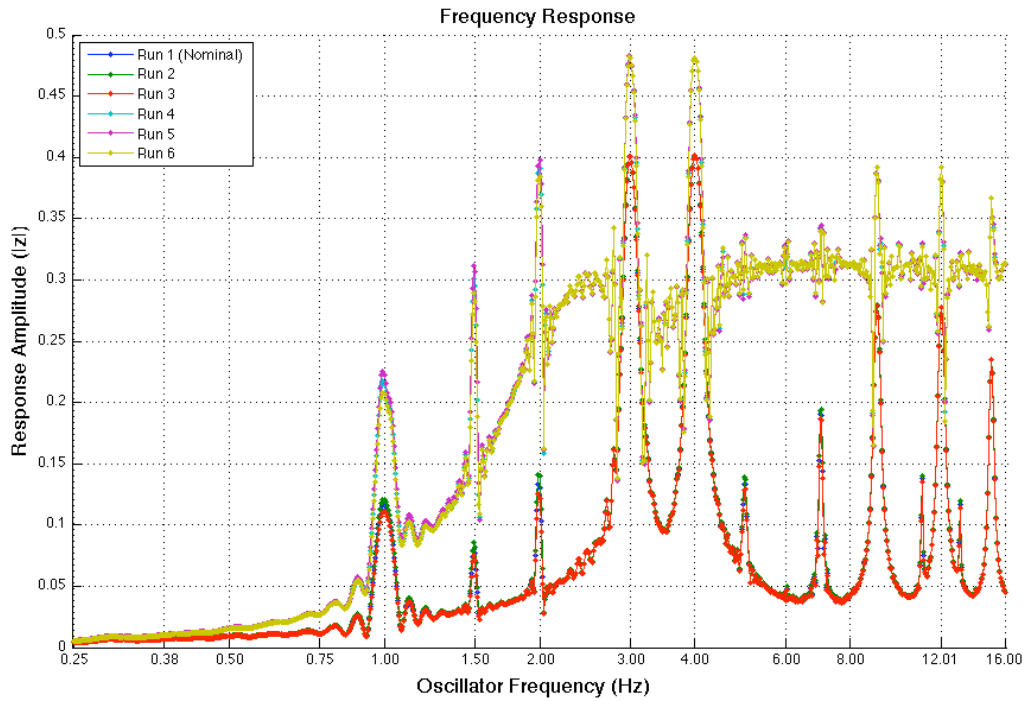


Figure 7.7: Frequency response patterns corresponding to sensitivity analysis for bifurcation parameter ( $\alpha$ ) and nonlinearity parameter ( $\epsilon$ ). The stimulus is a 4:3 polyrhythmic pattern with 1 Hz repetition frequency. The numerical values of the parameters of each run are listed in Table 2.

It seems like all six runs can be reduced into two distinctive patterns of responses. In fact, by calculating the coefficient of similarity between pairs of runs we get that run 1 (nominal values) and run 2 are highly similar with  $r^2 = 0.998$ . Similarly, run 1 is highly similar to run 3 with  $r^2 = 0.998$ . Run 2 and run 3 are also similar with  $r^2 = 0.992$ . With these degrees of similarity among the first three runs we can consider the nominal run as representative of runs 1, 2 and 3. In the first three runs only the value of parameter ( $\epsilon$ ) is altered and we can observe that such numerical manipulation within the described range has virtually no effect in altering the frequency response pattern, thus the neural resonance model produces responses that are resilient with regard to numerical manipulation of parameter ( $\epsilon$ ) exclusively.



Similarly, calculation of the  $r^2$  index between run 4 and run 5, and between run 5 and run 6 gives 0.998 for both cases. Run 5 and run 6 have  $r^2 = 0.995$ . Under these circumstances we can take run 4 as representative of runs 4, 5 and 6. The responses of the two groups of runs identified above are combined in Figure 7.8 below. The green graph represents the first group of runs (1-3) and the blue one represents the second group of runs (4-6). The coefficient of similarity for these two graphs is 0.4368, which indicates a weak degree of similarity. The differentiation in the frequency response pattern in the second group of runs in comparison to the first group is due to the 10% increase of the nominal value of parameter ( $\alpha$ ), because it has been shown that the numerical manipulation of parameter ( $\varepsilon$ ) has no impact in the behaviour of the neural resonance model.

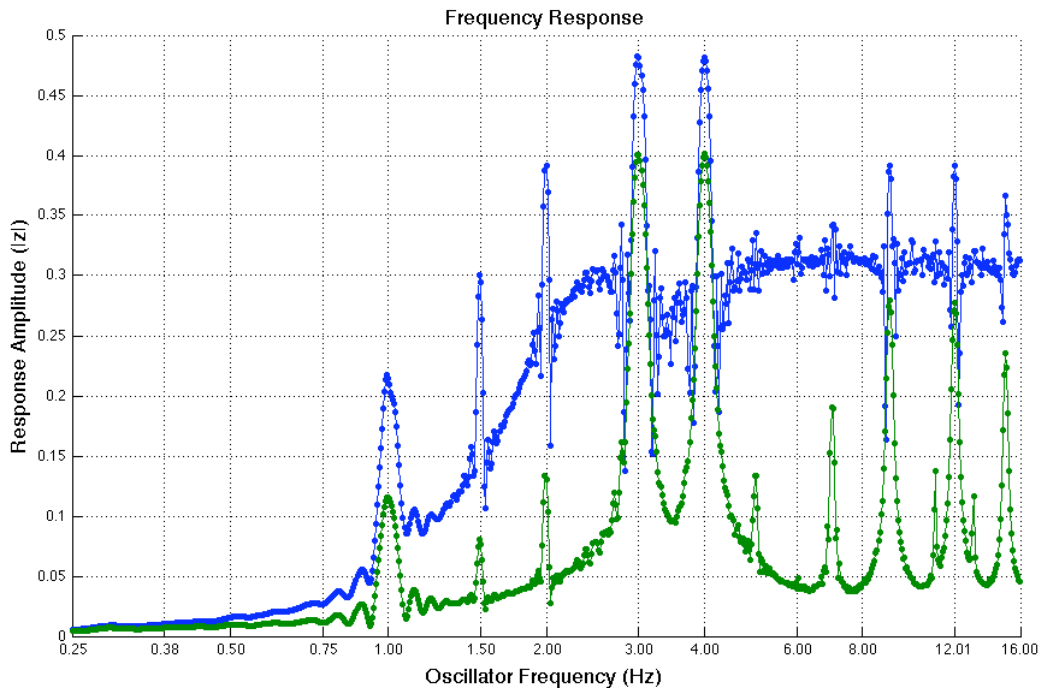


Figure 7.8: Frequency response patterns for ( $\alpha > 0$ ) (blue) and ( $\alpha = 0$ ) (green).

Despite the relatively low degree of similarity we can see that the resonance effect is similar in both patterns. In particular, in the blue graph we can still see that there are resonating oscillators with natural frequencies

similar to those observed in the green graph, such as 1Hz, 1.5Hz, 2Hz, 3Hz and 4Hz. The only difference is that the peak responses of the resonating oscillators is stronger when parameter ( $\alpha$ ) is small positive and also, the amplitude floor for those oscillators in the between of the resonating ones in the frequency range 1 Hz - 16 Hz is raised when ( $\alpha > 0$ ). However, in the case where a large negative value is given to parameter ( $\beta$ ) (the damping parameter) that prevents the oscillation amplitude from becoming large when ( $\alpha$ ) is small positive, then the two graphs are highly correlated with  $r^2$  equal to 0.9406 (not depicted here). From that perspective, it can be said that the frequency response patterns of the neural resonance model reported in Chapter 6 are consistent for a range of numerical values of parameters ( $\alpha$ ) and ( $\epsilon$ ).

#### 7.1.3.4 *Sensitivity analysis for parameters ( $\beta$ ) and ( $\epsilon$ )*

The set of simulations (runs) presented below assumes that parameter ( $\alpha$ ) is small positive throughout the different runs. This is because the main use of parameter ( $\beta$ ) is to prevent the amplitude of a spontaneous oscillation ( $\alpha > 0$ ) from becoming infinitely large. The sensitivity analysis is applied for parameters ( $\beta$ ) and ( $\epsilon$ ). In particular, both single and multiple parameter analysis is applied for parameters ( $\beta$ ) and ( $\epsilon$ ). The numerical range and the specific values for the simulations for parameters ( $\beta$ ) and ( $\epsilon$ ), are calculated according to the uniform approach. The total number of runs in this section is summarised in Table 7.3. The star next to the run number is given to avoid confusion with the runs described in Table 7.2.

Table 7.3: Combinations of runs regarding parameters ( $\beta$ ) and ( $\varepsilon$ ) considering both single (Run 2\* and Run 3\*) and multiple (Run 4\* to Run 6\*) parameter analysis. Range of values is calculated based on a uniform variation of  $\pm 10\%$  around nominal values.

Runs	Parameter ( $\beta$ )	Parameter ( $\varepsilon$ )	Color
1*	Nominal ( $-1$ )	Nominal ( $1$ )	Brown
2*	$+10\%(-0.9)$	Nominal ( $1$ )	Blue
3*	$-10\%(-1.1)$	Nominal ( $1$ )	Green
4*	$+10\%(-0.9)$	$-10\%(0.9)$	Red
5*	$-10\%(-1.1)$	$+10\%(1.1)$	Light Blue
6*	$+10\%(-0.9)$	$+10\%(1.1)$	Purple
7*	$-10\%(-1.1)$	$-10\%(0.9)$	Golden

Figure 7.9 illustrates the frequency response pattern of the neural resonance model for the seven runs described in Table 7.3 above in the presence of a 4:3 polyrhythmic stimulus. The frequency response patterns exhibit a high degree of similarity among them. In this set of runs parameter ( $\alpha$ ) has been kept fixed ( $\alpha = 0.1$ ). In fact the resonance pattern is similar to the one produced by runs 4, 5, and 6 in the first set of runs (Table 7.2), where ( $\alpha$ ) was small positive. There we have seen that despite the small similarity suggested by the coefficient  $r^2$  between the graphs corresponding to ( $\alpha = 0$ ) and ( $\alpha = 0.1$ ), the fact that the same oscillators resonate across different runs suggests that the model's behaviour is not influenced by any small numerical manipulation of the parameters.

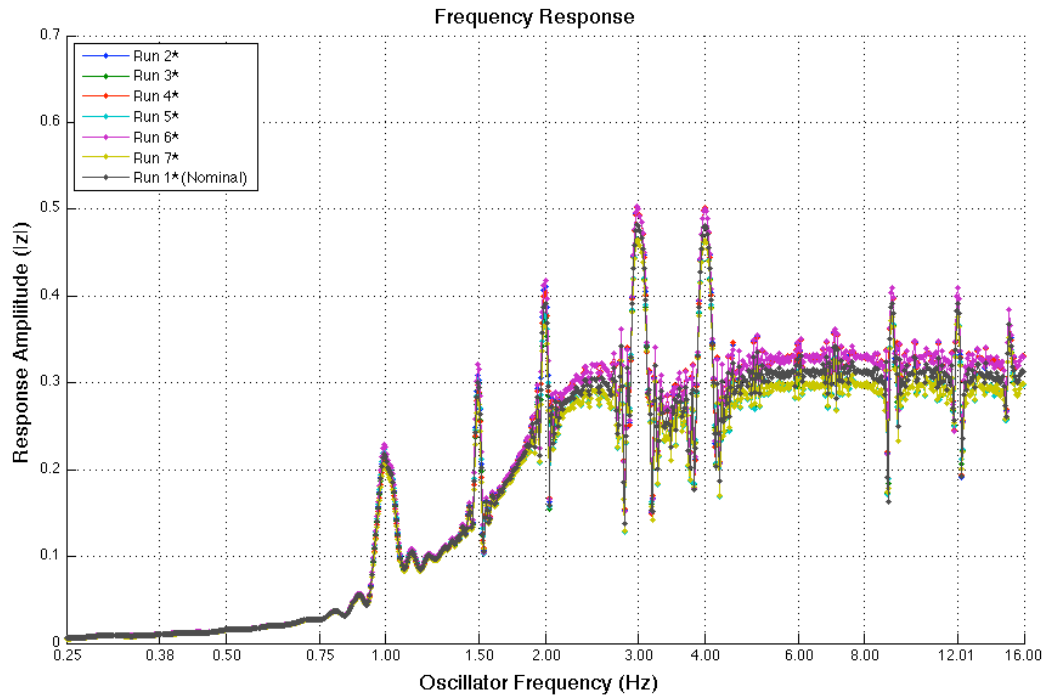


Figure 7.9: Frequency response patterns corresponding to sensitivity analysis for saturation parameter ( $\beta$ ) and nonlinearity parameter ( $\varepsilon$ ). The dynamic system is on spontaneous oscillation mode ( $\alpha > 0$ ). The stimulus is a 4:3 polyrhythmic pattern with 1 Hz repetition frequency. The numerical values of the parameters of each run are listed in Table 7.3.

So far we have employed parameter sensitivity analysis for a range of parameters related to the state variable of the neural resonance model. In general, we could say that parameters ( $\beta$ ) and ( $\varepsilon$ ) make almost no difference in the frequency response pattern, a positive value for parameter ( $\alpha$ ) raises the floor while a negative value leads to suppression of the ability of the system to resonate. The frequency response pattern of each of the simulations in this subsection has been suggestive of a model that produces consistent results despite any numerical manipulation of its parameters within a justified range, and therefore, it can be argued that the results presented in Chapter 6 are robust and not significantly dependent on the chosen numerical values of the parameters.

## 7.2 CROSS-MODEL COMPARISON WITH THE WILSON-COWAN MODEL

In the introduction to this chapter, we noted that false assumptions regarding the actual biological processes to be modelled can introduce uncertainty in the adequacy of modelled data for making comparisons with empirical data. As argued in the introduction, one way to deal with this uncertainty is to consider alternative similar-purpose models, ideally models which are considered to be particularly accurate or the best available in some relevant respect, regardless of their complexity, and to compare the behaviours of the two models (Haefner, 2005).

As indicated earlier, our first step in such a process has therefore been to choose an alternative similar purpose model to the neural resonance model. The model we have chosen to use here is the Wilson-Cowan model, which describes the dynamics of interaction between excitatory and inhibitory neurons. In fact, the neural resonance model derives from the Wilson-Cowan model, therefore we could say that in many respects the former shares indirectly the same biological assumptions as the Wilson-Cowan model. In Chapter 4 the general view was noted of the Wilson-Cowan model as a highly biological model, and therefore it can be considered as a model that limits uncertainty regarding biological assumptions more than any other obvious candidate model.

The second step in the process of cross-model comparison is to stimulate both models using the same polyrhythmic stimulus and to compare their behaviours. In cases where these behaviours are qualitatively similar it can be argued that the neural resonance model can be used interchangeably with the Wilson-Cowan, thus, any potential uncertainty with respect to the biological assumptions of the neural resonance model is limited up to the point that the assumptions of the Wilson-Cowan model are rigorous.

The Wilson-Cowan model simulates the dynamics of a neural oscillator, i.e., the properties of the interaction of two types of neurons (excitatory and inhibitory) that comprise a population of neurons (neural oscillator), with parameters corresponding to physical aspects of such a system. A detailed description of the equation of a Wilson-Cowan oscillator can be found in Campbell and Wang (1996) paper 'Synchronisation and desynchronisation in a network of locally coupled Wilson-Cowan (WC) oscillators'. We will not duplicate that description here. However, it will be useful for our purposes to consider aspects of the parameters used in these equations. In running simulations with the Wilson-Cowan model, values must be assigned to the parameters of the WC oscillators, just as in the case of the neural resonance model. Some of the parameters of a WC oscillator are determined externally. These parameters are:

- a) a parameter for self excitation in the excitatory neuron,
- b) a parameter that denotes the strength of the coupling from the inhibitory unit to the excitatory one,
- c) a parameter of the natural frequency of the oscillator, and
- d) two parameters reflecting the threshold of activation for both excitatory and inhibitory neurons.

Also, there are two parameters, whose values are calculated internally based on a set of nominal values given in the previous parameters. These are:

- a) a parameter for self excitation of the whole unit (neural oscillator), and
- b) a parameter that denotes the strength of the coupling from the excitatory unit to the inhibitory one.

In the following simulations we have been using a nominal set of values taken from the literature (Large et al., 2010) for the Wilson-Cowan model that is being used as a standard of comparison. It has been beyond the scope of this thesis to recursively apply parameter sensitivity analysis to the Wilson-Cowan model itself. As a result, we have to keep in mind that the discussion below is based on the responses produced given a particular nominal set and that for different values it is possible that the observed responses would have been different.

#### *7.2.1 Frequency response pattern of the Wilson-Cowan model to a 4:3 polyrhythmic stimulus with 1 Hz repetition frequency.*

Figure 7.10 illustrates the frequency response pattern of the Wilson-Cowan model when stimulated with a 4:3 polyrhythmic stimulus for a range of different stimulus's amplitudes. The use of different amplitudes is used here mainly to illustrate the frequency detuning property of the Wilson-Cowan model. The pattern is qualitatively similar to the neural resonance model, e.g., the resonating oscillators reflect the general resonance pattern suggested in Chapter 6 (Table 6.2). Also, similar to the neural resonance model, we can see that as the stimulus amplitude increases, the higher-order resonances at harmonic, subharmonic and more complex ratios with respect to the fundamental frequencies of the stimulus become more prominent. For example if we concentrate on the light blue trace we can see that for the 4:3 polyrhythmic stimulus, the responses we get correspond to those of the neural resonance model. In Figure 7.13 responses of both models are combined to illustrate this, but before that we need to elaborate on the observed frequency detuning in the frequency

response pattern of the Wilson - Cowan model. Figure 7.10 suggests that detuning increases in line with an increase in the stimulus amplitude.

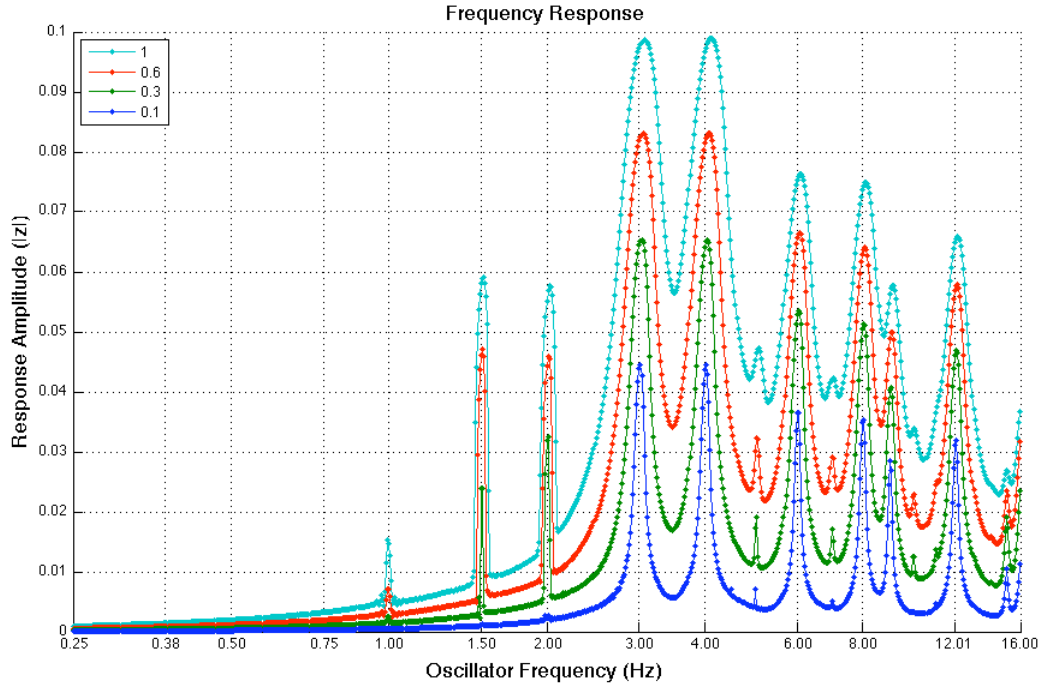


Figure 7.10: Frequency response patterns of the Wilson-Cowan model in the presence of a 4:3 polyrhythmic stimulus with 1 Hz repetition frequency for a range of different amplitudes of the stimulus.

### 7.2.2 Frequency detuning in the neural resonance model

Earlier in this chapter we have mentioned that frequency-detuning parameters are also present in the neural resonance model, i.e., parameters  $(\delta_1)$  and  $(\delta_2)$ . In the simulations of Chapter 6 no detuning has been observed in the frequency response patterns because the parameter related to detuning had taken a nominal value of 0, as disclosed in this chapter, and therefore was inactive. However, by changing the value of parameter  $(\delta)$  to  $-9$ , which is the nominal value used by Large, Almonte, and Velasco (2010) we can expect detuning to arise in the frequency response pattern



of the neural resonance model. This is illustrated in Figure 7.11 below. We can see that the peaks are slightly bent to the right as a result of a negative value of ( $\delta$ ).

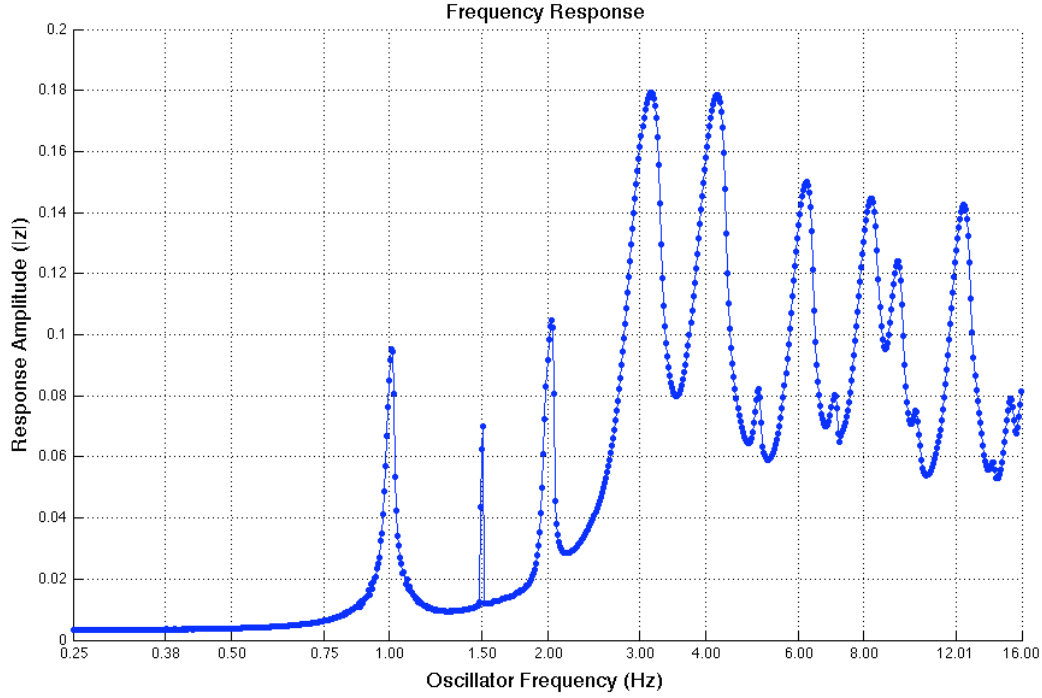


Figure 7.11: Frequency response pattern of the neural resonance model in the presence of a 4:3 polyrhythmic stimulus with 1 Hz repetition frequency for a negative value of the frequency detuning parameter ( $\delta$ ).

### 7.2.3 *Superimposed frequency response patterns for both the neural resonance and the Wilson-Cowan model*

Finally, by superimposing the frequency response patterns of both models in Figure 7.12 below we can observe the structural similarity between the two responses. The coefficient of similarity is  $r^2 = 0.8677$ .

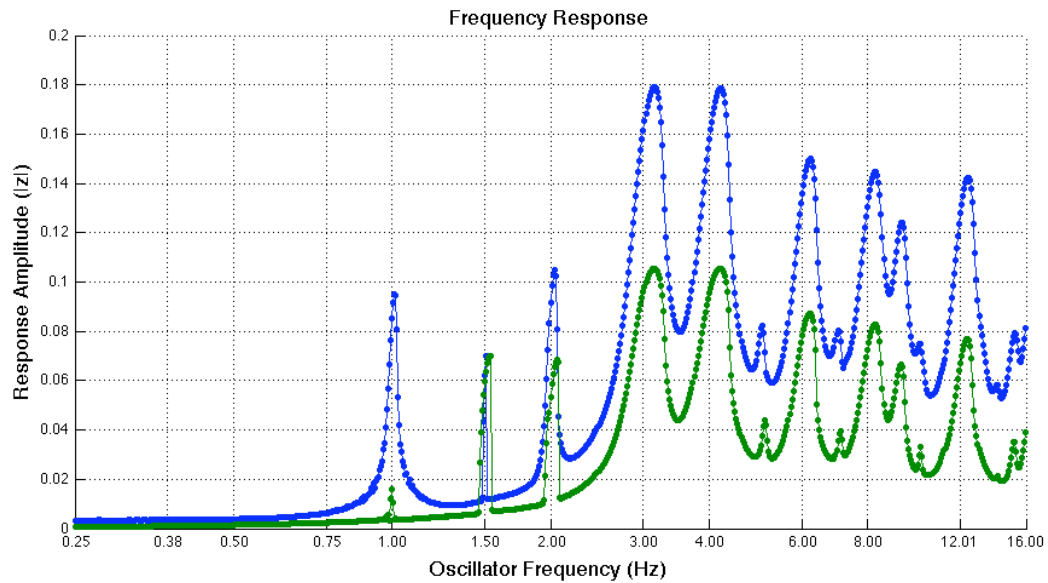


Figure 7.12: Superimposed frequency response patterns of Wilson-Cowan and neural resonance model. The two patterns are quite similar with  $r^2 = 0.8677$ .

### 7.3 SUMMARY

In this chapter we have applied systematic techniques to limit the uncertainty that arises in relation to reliance on data from the neural resonance model. In particular, two such techniques were employed: parameter sensitivity analysis and cross-model comparison. Firstly, in the case of parameter sensitivity analysis, systematic numerical manipulation was applied to the parameters of the state variable, the parameters relating to the driving variable, and to the initial condition of the amplitude of the oscillators. This allowed us to rule out dependency of the results presented in Chapter 6 on specific numerical values.

Secondly, we compared the behaviour of the neural resonance model with the behaviour of the Wilson-Cowan model in order to limit uncertainty with regard to the biological assumptions of the neural resonance model. The Wilson-Cowan model was chosen as the model

that best limits uncertainty of biological assumptions more than any other candidate model. Simulations using the Wilson-Cowan were found to produce outputs qualitatively the same as those of the neural resonance model. However, for completeness we should note that certain nominal values were used to run the simulations with the Wilson-Cowan model without recursively applying sensitivity analysis to the parameter of the Wilson-Cowan model.

The neural resonance model derives from the Wilson-Cowan model, and shares, in less explicit, but more mathematically convenient, form its biological assumptions. Given that the Wilson-Cowan model is generally considered (Chapter 4) to be the most biologically faithful of the candidate models, and bearing in mind the similarity of responses between the two models, we are able to place a reasonable limit on uncertainty with regard to the biological assumptions made by the neural resonance model.

## MODEL PREDICTIONS ABOUT HUMAN BEHAVIOUR

---

In this chapter the neural resonance model is used to make predictions about human tapping behaviour in polyrhythm perception, which are then compared with empirical data collected by conducting novel experiments with people. In the next few sections we will focus in detail on two as-yet-unexamined predictions made by the neural resonance model about human tapping behaviour in association with polyrhythm perception. These predictions are as follows.

1. The categorisation used by (Handel and Oshinsky, 1981) to summarize observed human tapping behaviours in polyrhythm perception is incomplete - new kinds of tapping are predicted based on the behaviour of the model (restated below).

2. The time it takes for an oscillator to reach relaxation time makes predictions about the relative times humans will take to start tapping in different ways to a polyrhythm. Our specific predictions are based on a refined version of a hypothesis suggested by (Large, 2000). Large previously made some limited tests of the hypothesis (as discussed below).

We then discuss how the predictions inform the design of the experimental study and present the data obtained from doing the actual experiment. We conclude by evaluating the model's predictions by comparing modelled data with empirical data. We will see that the two principal results are as follows. Firstly, there is an empirical confirmation of a prediction by the model of a category of human tapping behaviour not

noticed in previous empirical studies. Secondly, the time it takes for people to start tapping along with a polyrhythm - as soon as they feel a beat - shares the same timescale with the temporal development of the resonance dynamics in the network for reaching a steady state.

### 8.1 PREDICTIONS TO EXTEND THE CATEGORIES OF HUMAN TAPPING BEHAVIOURS ASSOCIATED WITH POLYRHYTHM PERCEPTION

The essence of the first of the two new predictions is simple. Namely the frequency response pattern of the neural resonance model predicts that the categories of human tapping behaviours observed in the empirical study of Handel and Oshinsky (1981) is not exhaustive, and new kinds of tapping are predicted. In Chapter 2, we saw that for a polyrhythm, which consists of two pulse trains (m) and (n), the categories of observed tapping behaviours reported by the Handel and Oshinsky study expressed in rhythmic frequencies were as follows:

1. Tapping along with the fundamental frequency of (m).
2. Tapping along with the fundamental frequency of (n).
3. Tapping along with the 1st subharmonic of (m).
4. Tapping along with the 1st subharmonic of (n).
5. Tapping along with the repetition frequency of the polyrhythmic pattern.

In Chapter 6, we saw that the neural resonance model not only captures all the aforementioned frequencies when stimulated by a n:m polyrhythm for a range of tempi, but also, it produces several other frequency responses that include:

1. the 2nd harmonic of (n),
2. the 2nd harmonic of (m),
3. the 3rd harmonic of (n),
4. the 3rd harmonic of (m), and
5. the 1st and 2nd subharmonic of the pattern's repetition frequency.

Subject to one important proviso, the above components of the frequency response pattern of the model may be taken as predictions for observable human tapping behaviours in polyrhythm perception. The important proviso, and the reason for specifically concentrating on the particular components noted above and not others from the full range one can observe in a frequency response pattern was discussed in Chapters 5 and 6, is as follows. Namely, some, but not all of these additional components correspond to humanly attainable tapping frequencies, because their underlying rhythmic frequencies fall within the range of upper and lower limits reported in studies of sensorimotor synchronization tasks (Repp, 2005). Given that important proviso, in order to provide empirical evidence in support to the above predictions, we decided to undertake a study similar to, but more comprehensive than, that of Handel and Oshinsky.

## 8.2 PREDICTION OF THE TIMESCALE PEOPLE START TO TAP TO POLYRHYTHMS

The second prediction is based on modifying a hypothesis suggested and subjected to limited tests by Large (2000). According to the original hypothesis, the time it takes for humans to start tapping along with a given stimulus - as soon as they feel a beat - can be associated with the time it takes for oscillators to reach relaxation time, i.e., a peaked and

approximately stabilized amplitude response (Large, 2000). In this thesis, we adopt a modified version of Large's hypothesis due to the following two reasons. Firstly, we argue that the method Large uses to test his hypothesis (in particular the summary statistics he is referring to) needs to be revised for concrete reasons that will be described in the text. Secondly, we propose an alternative way of measuring the relaxation time of the oscillators' resonances focusing on the timescale of the relaxation as opposed to an absolute point in time where relaxation occurs. In general, relaxation time corresponds to the time it takes for a perturbed system to return into equilibrium. As a scientific convention, the simplest theoretical description of relaxation as function of time  $t$  is an exponential law  $\exp(-t/\tau)$ , where  $(\tau)$  is the relaxation time. In our case relaxation time refers to the time it takes for an oscillator to reach a stable resonance pattern in the presence of a stimulus, that is the time it takes for an oscillator to reach its maxima magnitudes around which it stabilises, however some small fluctuation is allowed.

### 8.2.1 *Measuring relaxation time*

In order to measure relaxation time, we can refer to graphs that show the development of resonances, i.e., the way the amplitude of an oscillator increases over time, for all the oscillators in the bank. Figure 8.1 below shows the development of such dynamics in the presence of a 4:3 polyrhythmic stimulus with 1 Hz repetition frequency. The x-axis corresponds to time in seconds, the y-axis corresponds to frequencies of the oscillators in the bank (in this case 768 oscillators from 0.25 Hz to 16 Hz as discussed in Chapter 5). The color code represents the amplitude of the oscillator over time. We can see that at the beginning of the stimulation (0

seconds) the amplitude of all the oscillators is close to zero. This is related to the chosen initial condition of the amplitude of all the oscillators in the bank (see Chapter 7).

We can see that oscillators with natural frequencies around the input frequencies 3Hz and 4Hz of the pulse trains comprising the polyrhythm start to resonate straight after the 1<sup>st</sup> second of the stimulation period, and within the next three seconds of the stimulation period they attain a quasi-stable prominent amplitude. The time to reach this amplitude is the relaxation time. The band of oscillators with natural frequencies that are subharmonically related to the stimulus input frequencies, start to resonate in a delayed manner in comparison to the previous set of oscillators, and also, they need more time to relax. In particular, these are the oscillators with frequencies 2Hz and 1Hz. They also start resonating from the 2<sup>nd</sup> second onwards but in a not so strong way in comparison to the previous set of oscillators. Furthermore, they need more time to relax in comparison to the previous set of oscillators. In addition to the previous oscillators, there are two oscillators that are harmonically related (in fact corresponding to the 2<sup>nd</sup> harmonics of the fundamental frequencies of the pulse trains) to the stimulus input frequencies and exhibit similar behaviour to the first group of oscillators described above. These are the oscillators with 6Hz and 8Hz frequencies. Other oscillators also resonate in the presence of the 4:3 polyrhythm and are all related to the general framework of resonances described in Chapter 6 (see Table 6.2).



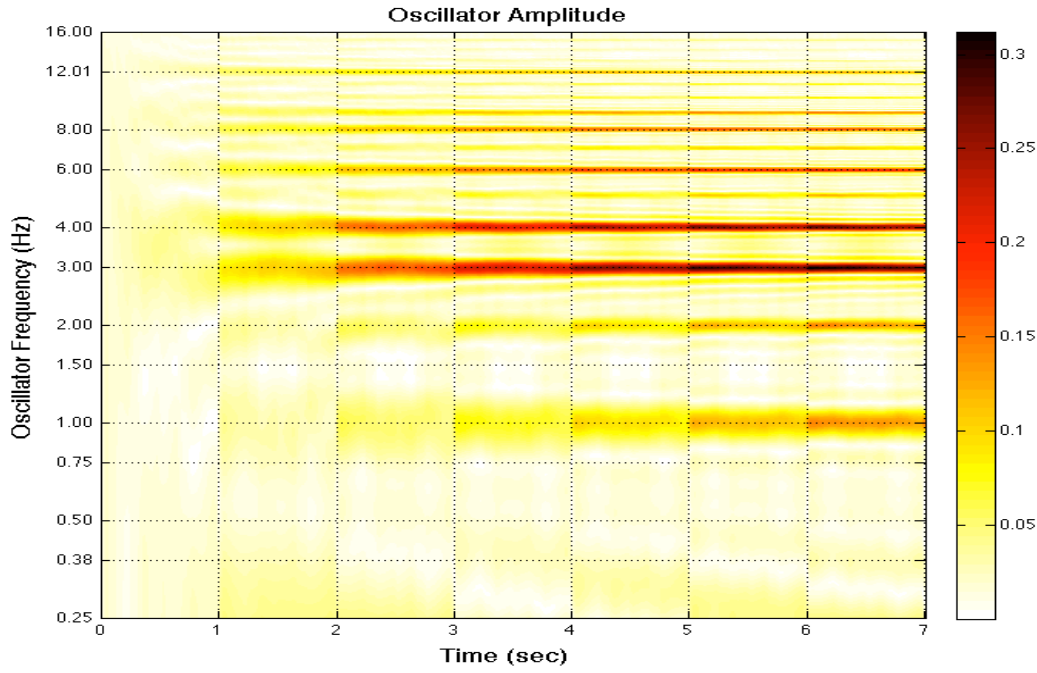


Figure 8.1: Amplitude responses of the bank of oscillators in the presence of a 4:3 polyrhythmic stimulus with 1 Hz repetition frequency.

The above amplitude responses in the presence of a polyrhythmic stimulus in combination with the hypothesis that the time it takes for oscillators to relax may be indicative of the time it takes for people to start tapping (Large, 2000) can be used to make testable predictions about the time it takes for humans to perceive beat in polyrhythms.

### 8.2.2 *Modifying the original hypothesis*

We modify Large's hypothesis, by focusing on the development of the dynamics of each oscillator over time, to make predictions about the timescale within which people start to tap along with a polyrhythm as soon as they feel a beat, instead of focusing on the absolute point in time when the oscillators relax. In other words, our modified hypothesis focuses on

the timescale of the dynamics development rather than exact points in time, such as when relaxation is reached. This allows us to make predictions about the relative time it takes for people to start tapping.

### 8.2.3 *Method limitations in evaluating the original hypothesis*

The way Large (2000) evaluates his hypothesis is by comparing the absolute relaxation time of an oscillator with the average time it takes for people to start tapping. However, this method may not be the most appropriate, given that our experimental observations suggest that the distribution of the time to start tapping is not a normal distribution, thus the use of mean as a measure of location may be biased.

Also in order to make his comparisons, Large (2000) uses the mean time associated with one particular empirical dataset among the three different datasets available in the empirical study he refers to. In other words, it seems like a specific empirical result is paired to match the behaviour of the model.

## 8.3 EXPERIMENTAL STUDIES IN HUMAN PERCEPTION OF POLYRHYTHMS

According to the predictions described in the previous section, the aim of our new proposed experiments with people is to monitor two main variables. These are the time it takes for people to start tapping along with a polyrhythmic stimulus and the category of the tapping behaviour. In Chapter 5 we have described the methodology for obtaining the above type of empirical data.

### 8.3.1 *Experimental design and participants*

The experiment is concerned with obtaining data for all five types of polyrhythmic stimuli and for all ten different repetition rates employed in the Handel and Oshinsky study. As a result, 50 combinations were presented randomly to each one of the 37 participants of the experiment, which gave rise to 1850 trials in total. 18 female and 19 male participants were recruited randomly across the Open University campus. The mean age of the group was 40 years old with standard deviation 11 years. 10 participants (27%) had 0 years of formal musical training, 12 participants (32%) had between 1 and 5 (inclusive) years of training, and the remaining 15 participants (41%) stated formal training experience between 5 and 34 years. Almost all the participants (92%) stated they are music enthusiasts, i.e., selected either 4 or 5 to a 5-point scale with 5 representing the most enthusiastic.

### 8.3.2 *Data management and visualisation*

Below we present an example that illustrates the process of pooling raw data, sorting the 1850 trials into classes, analysing each trial within each class to identify the category of tapping behaviour (by means of finding the statistical median of the distribution of the values of the inter-tap intervals), creating summary datasets for each class that capture both the category of tapping and the time it takes to start tapping, and visualizing these datasets by using boxplots. In effect, the total number of summaries at the end of the analysis would be the same to the number of combinations presented to people, i.e., 50. Each summary contains a number of columns that corresponds to the number of the categories of tapping behaviours

observed, and numerical entries that correspond to the time it takes for people to start tapping each one associated with a particular category. In this chapter we present 10 summaries regarding the continuing example of a 4:3 polyrhythm repeating at all different tempi, and a further two summaries for a 3:2 and a 2:5 polyrhythm repeating once per second as a sufficient number of summaries to communicate the argument of the chapter.

### 8.3.3 *Raw data*

Figure 8.2 is an excerpt of the spreadsheet where raw data were imported for initial sorting. It contains six randomly selected complete trials among three participants. In particular, there are three complete trials from participant 8, two trials from participant 9, and one trial from participant 10. Each trial can be distinguished by its ID on the top of the list, and the rest of the numerical entries correspond to a point in time where a tap has occurred in msec after the participants have started listening to the polyrhythmic stimulus. The ID of that trial is denoted for example by the following sequence of numbers 1/ 0. 3. 4. 0. The first number (1) corresponds to the repetition frequency of the polyrhythm, which in this case is 1 Hz. The remaining sequence of the numbers (after the forward slash) is used to denote the type of polyrhythm presented to the participant, which in that case is a 4:3 polyrhythm. The first numerical entry below the ID for that particular trial is 4004 and implies that the first tap in the trial was recorded 4004 msec after the polyrhythm was first heard. In that trial, participant 8 has decided to finish the trial after tapping for 8 times along with the polyrhythmic stimulus.

participant8	participant9	participant10
1/ 0. 3. 4. 0	0.555556/ 0. 3. 0. 5	0.833333/ 2. 3. 0. 0
4004	3562	2052
5018	4082	2806
6006	4576	3560
7046	5070	4468
8060	5590	5222
9022	6032	6002
10010	6604	6834
10998	7122	7614
0.5/ 2. 3. 0. 0	7642	8392
3980	8164	9222
4656	8760	10002
5306	9306	10782
5956	0.625/ 2. 3. 0. 0	11640
6632	6420	12446
7334	7016	13200
7984	7562	14032
8634	8108	14862
9284	8628	15642
1.666667/ 2. 3. 0. 0	9148	16446
2442	9642	17226
3040	10188	18032
3610	10708	18862
4234	11228	19668
4832	11774	20448
5428	12346	21252
6000	12840	22032
6600	13386	22812
	13932	23644

Figure 8.2: Excerpt of pooled raw empirical data

So far, the raw data -in total- contains timing information on tapping behaviours for the total number of 1850 trials. Next, we discuss how to identify which tapping category has been chosen by the participant and whether that category complied with a regular pulse-like tapping as required for the purposes of the experiment. For example, in some trials some participants chose to shadow (copy) parts or all of the polyrhythmic pattern despite the instruction of trying to refrain from doing this, or participants changed the mode of tapping during the course of the trial, or they gradually drifted out of phase with respect to the targeted beat, which eventually led the participant to abandon the trial. In such cases the trials were dismissed from further analysis.

#### 8.3.4 *Identifying the category of tapping behaviour: an example*

A single trial from those illustrated above can be used as an example to explain how the category of tapping behaviour for each individual trial is identified. The number of numerical entries correspond to points in time where successive tapping behaviours occur, therefore we can calculate the inter-tap interval (ITI) in msec, which is the time that elapses between two successive taps. Figure 8.3 illustrates all the ITIs associated with the first trial of participant 8. In order to identify the chosen category of tapping we calculated the median of the ITI entries. We have selected to calculate the median ITI as a more appropriate statistical summary than the average ITI because the former is less susceptible to outlying observations caused for example by participants significantly drifting or losing 'steadiness' in their tapping. The median is calculated by putting the numerical entries in ascending order and selecting the middle value in the range. In case where the total number of entries is an odd number then the median is the single number in the middle of the range. If the total number of entries is an even number then the median is the average of the two values in the middle of the range.

	Timing (msec)	Inter-tap Interval (msec)
1st tap	4004	1014
2nd tap	5018	988
3rd tap	6006	1040
4th tap	7046	1014
5th tap	8060	962
6th tap	9022	988
7th tap	10010	988
8th tap	10998	Median (msec)
		988

Figure 8.3: Timing data for successive taps in the presence of a 4:3 polyrhythmic stimulus repeating once every 1000 msec. The third column corresponds to inter-tap intervals in msec for successive pairs of taps (ITIs). The median is also calculated in the bottom left-hand side of the third column and is used to identify the mode of tapping chosen by the participant. In this trial the participant taps approximately every 1000 msec, therefore the mode of tapping is along with the co-occurrence of the two pulse trains.

In the above trial, the participant starts tapping after 4004 msec, which means that he starts tapping for the first time at the beginning of the 5th repetition of the 4:3 polyrhythmic pattern. The median of the ITIs is approximately 1000 msec, therefore the participant chooses to tap along with the co-occurrence of the two pulse trains. Over the course of the trial there is a small deviation from perfect periodicity in the tapping behaviour, which is commonly identified in empirical studies where people are asked to tap in a regular way. In fact, empirical studies have shown that for a musically trained and practiced individual at least 68% percent of his asynchronies, i.e., variability from perfect periodic tapping, will fall within a range of 1 standard deviation, that is the 2% of the IOI (Repp, 2000). For example, in the above case where the target is to tap along with the co-occurrence of the two pulse trains, i.e., once every 1000 msec, a standard deviation of 2% of the IOI would be  $1000 * 0.02 = 20$  msec. According to Repp's suggestion approximately 68% of the asynchronies in this trial would fall within a range from 980 msec to 1020 msec. Indeed, 5 out of 7 of the above ITIs (approximately 71%) fall within the aforementioned

range. This result implies a trained and practiced individual according to Repp's findings. Interestingly, in the personal data form filled prior to the experiment participant 8 reported 15 years of formal musical training and 45 years of practicing experience. However, a collective analysis for all the 50 trials undertaken by that participant would be needed to provide a more accurate view on the participant's sensorimotor synchronization skills.

In any case, for trials with similar characteristics to this, such as small deviation around the targeted beat and consistency in keeping the same beat, the use of the median to identify the category of tapping is meaningful. The above process of calculating the median and as a result identifying the category of tapping behaviour was repeated for all 1850 trials. Some of the trials were dismissed from further analysis on the basis described in the last paragraph of the previous section. An example of a dismissed trial is illustrated below (Figure 8.4). The presented polyrhythm is of a 4:3 type repeating once every 600 msec.

	Timing (msec)	Inter-tap Interval (msec)
1st tap	7046	780
2nd tap	7826	468
3rd tap	8294	1092
4th tap	9386	442
5th tap	9828	858
6th tap	10686	390
7th tap	11076	806
8th tap	11882	364
9th tap	12246	806
10th tap	13052	364
11th tap	13416	884
12th tap	14300	338
13th tap	14638	754
14th tap	15392	390
15th tap	15782	

Figure 8.4: An example of a dismissed trial.



Based on the first three numerical entries of ITIs in column 3 of Table 8.3 we may infer that the participant starts tapping without having an explicit pulse-like perception already formed in his mind. From the 4th ITI onwards we may further infer that the participant copies part of the polyrhythmic stimulus inasmuch there is an alteration of two ITIs of rough durations around 400 msec and 800 msec.

The outcome of the above processing of raw data (in terms of ITI and Median calculation) leaves a pool of trials that is less than 1850 in total, and for each one of those trials left in the usable pool we have two pieces of useful information that can eventually be compared with the neural resonance model predictions. In other words, for each trial the category of a tapping behaviour and the time it takes for a participant to start tapping in this way along with a polyrhythmic stimulus are both known. One convenient way to get an idea of how many trials are dismissed is to look at the collective results for each individual combination, i.e., polyrhythmic type and presentation rate, which is presented in the next section. For example, one particular combination would ideally contain 37 trials assuming that all the participants perform in line with the instructions of the experiment. However, it is expected that the frequency of successful trials varies across combinations. In the next section we give an example of creating a collective dataset that provides information about the categories of tapping behaviours and their corresponding times, i.e., the time it takes to start tapping, with regard to a specific combination, i.e., a 4:3 polyrhythm repeating once per second.

### 8.3.5 Creating a dataset for each combination, i.e., polyrhythm type and tempo

The total number of trials is organized into 50 datasets, each one representing a different combination. For each combination the number of columns of the dataset corresponds to the total number of observed categories of tapping behaviours. The numerical entries in each column are independent to each other and correspond to the time it takes to start tapping for different participants. Figure 8.5 illustrates one such dataset for a combination whereby the type of a polyrhythm is 4:3 and the repetition rate 1000 msec.

4-pulse train	3-pulse train	unit meter	1st sub of 4-pulse train	grouping 6 pulses of a 4-pulse
3170	2730	4004	5068	3042
3062	3976	7046	2080	
1338	4030	14294	2966	
	2392	5980	6056	
	3046	1143	3016	
	2054	3092	2158	
	7018	3068	3058	
	2056	12034	2028	
	1040		3016	
	3044		2054	
			2060	
			2992	
			2036	
			2007	
			4024	

Figure 8.5: Collective data for a single combination (in this case a 4:3 polyrhythm repeating once per second). Each column represents a single category of tapping and each numerical entry corresponds to the time it takes for a single participant to start tapping.

The adopted convention for describing the tapping categories is to express it in relation to the pulse trains of the polyrhythmic pattern. So in Figure 8.5, the first column corresponds to tapping along with the 4-pulse train and it is labeled as '4-pulse train', the second column corresponds to tapping along with the 3-pulse train mode and it is labeled as '3-pulse train', the third column corresponds to tapping along with co-occurrence of

the two pulse trains and by following Handel and Oshinsky's convention it is labeled as 'unit meter', the fourth column corresponds to tapping along with every other element of the 4-pulse train therefore it is labeled as '1st subharmonic of the 4-pulse train'. The last column corresponds to a tapping mode where the participant groups 6 pulses from the 4-pulse train and it is labeled as 'grouping 6 pulses of a 4-pulse'. A similar convention is employed for all 50 datasets. The numerical entries in each column correspond to the time it took for participants to start tapping in the presence of the polyrhythmic stimulus. The above dataset contains information about the categories of tapping behaviours along with a 4:3 polyrhythm, the time to start tapping, and also, preferences of categories in terms of the ratio number of numerical entries per category / total number of entries. Also, the total number of numerical entries is 37, which means that in this particular combination no trial has been dismissed. In the next section we introduce the use of boxplots as a convenient way to visualise concisely the dataset described in Figure 8.5.

### 8.3.6 *Visualisation of data using boxplots*

Considering the two experimental variables described above, i.e., category of tapping behaviour and time it takes to start tapping, the most appropriate graphical display/visualization of the obtained data is the use of multiple boxplots. Figure 8.6 illustrates a graph of multiple boxplots corresponding to the data presented in Figure 8.5.

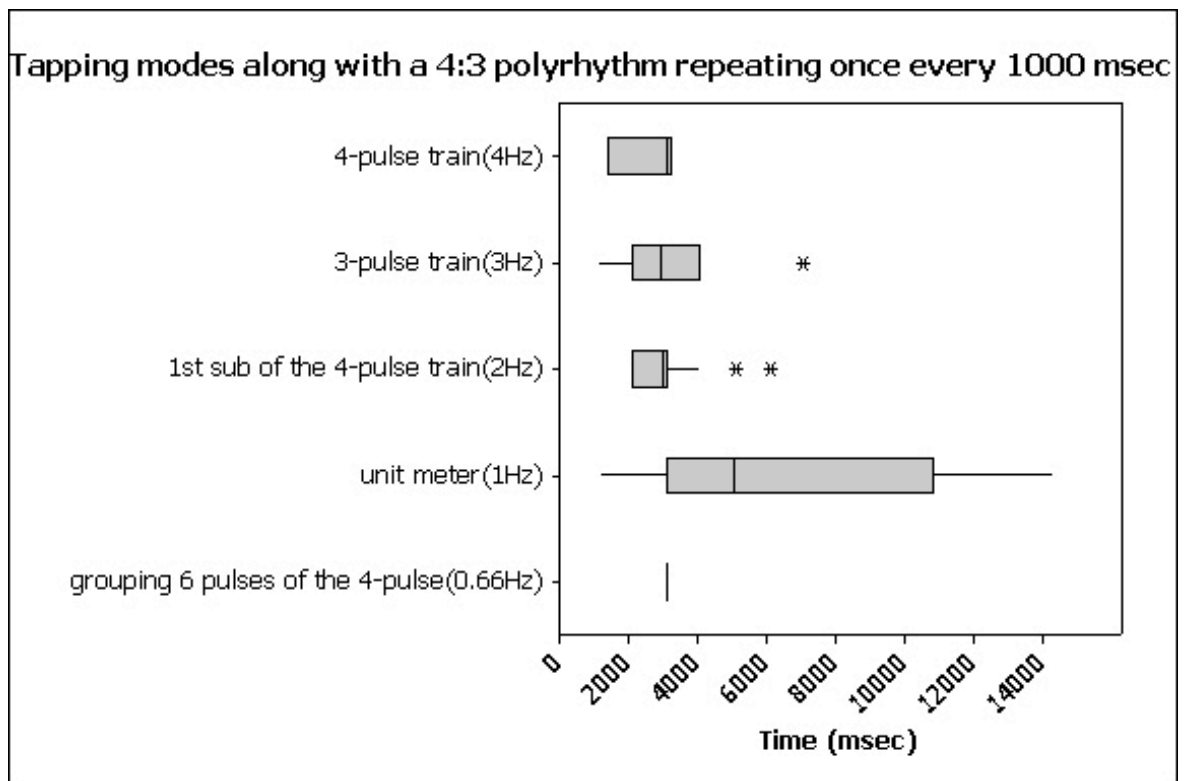


Figure 8.6: Multiple boxplot summarizing the categories of tapping and the range of time it takes for people to start tapping along with the given 4:3 polyrhythm repeating once every 1000 msec.

The x-axis corresponds to the time that elapses after the polyrhythm is first heard. On the y-axis the multiple categories of tapping behaviours are listed. For each category there is a boxplot associated with it. Each boxplot consists of the grey area box where 50% of all the timing observations fall. The right edge of the box corresponds to the upper quartile  $Q_3$  (75% of all observations) and the left edge to the lower quartile  $Q_1$  (25% percent of all observations). The line splitting the box is the median time it takes for people to start tapping. The difference between upper quartile and lower quartile is the inter-quartile range (IQR), which is also used to define the minimum and maximum values of the range, or lower and upper adjacent values respectively. These two values are furthest away from the median on either side of the box, but are still within a distance of 1.5 times the IQR

from the nearest end of the box, i.e., the nearest quartile. The lower and upper adjacent values are denoted by the end of the whiskers (the lines extending potentially from either ends of the box). The stars correspond to outliers, i.e. observations that do not fall in the range described above.

Figure 8.7 provides complementary details about the exact numerics regarding the multiple boxplot presented above. For example, it provides an accurate picture regarding the number of people who have chosen to tap in a certain ways. This information is useful for creating the frequency distribution of the categories of tapping behaviour.

Variable	Count	Minimum	Q1	Median	Q3	Maximum	IQR
4-pulse train	3	1338	1338	3062	3170	3170	1832
3-pulse train	10	1040	2056	2887	3990	7018	1934
1st sub of 4-pulse train	15	2007	2054	2966	3058	6056	1004
unit meter	8	1143	3074	4992	10787	14294	7713
grouping 6 pulses of a 4-pulse train	1	3042		3042		3042	

Figure 8.7: Descriptive statistics for the 4:3 polyrhythm repeating once per second.

The visualization of tapping data using boxplots in the way described above allows us to access information with regard to the two central experimental variables of the empirical tests, namely the categories of tapping behaviours and their corresponding time it takes for people to start tapping both in terms of range and median. In the last part of this section we list graphs of multiple boxplots for the remaining combinations with respect to a 4:3 polyrhythm repeating at the remainder 9 tempi. Multiple boxplot graphs regarding the 3:2 and 5:2 polyrhythm are presented in the last section of the chapter.

## 8.3.7 Empirical results of tapping along with a 4:3 polyrhythm for different tempi

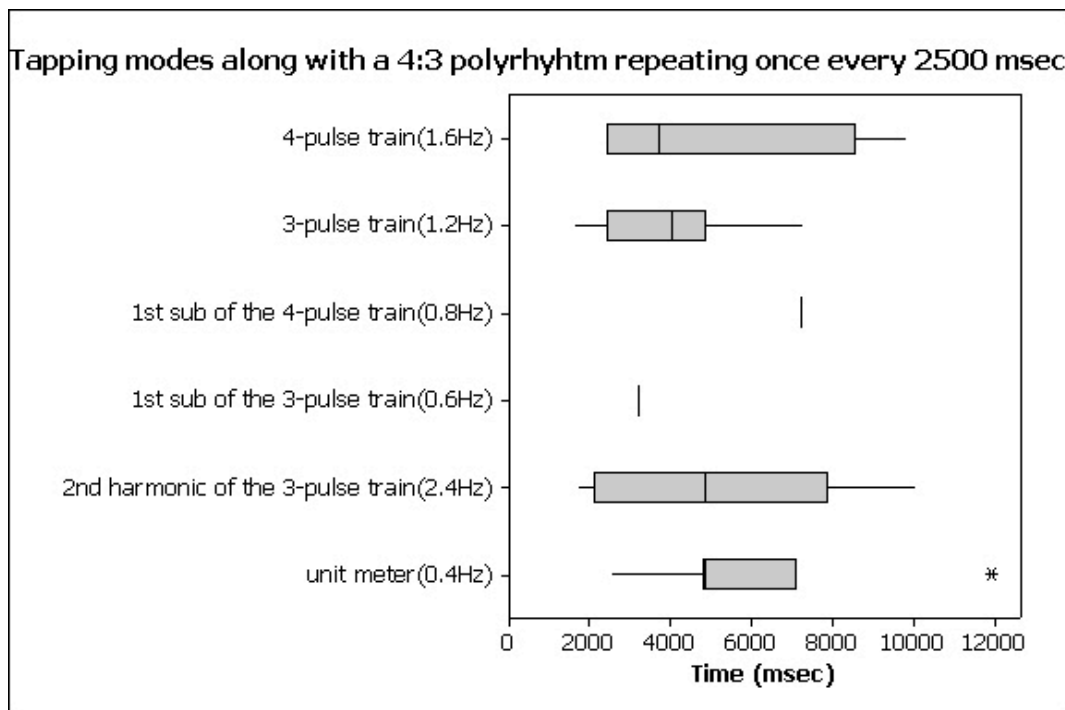


Figure 8.8: Multiple boxplots for the range of tapping modes and their corresponding period of time within which people start tapping along with a 4:3 polyrhythm repeating once every 2500 msec.

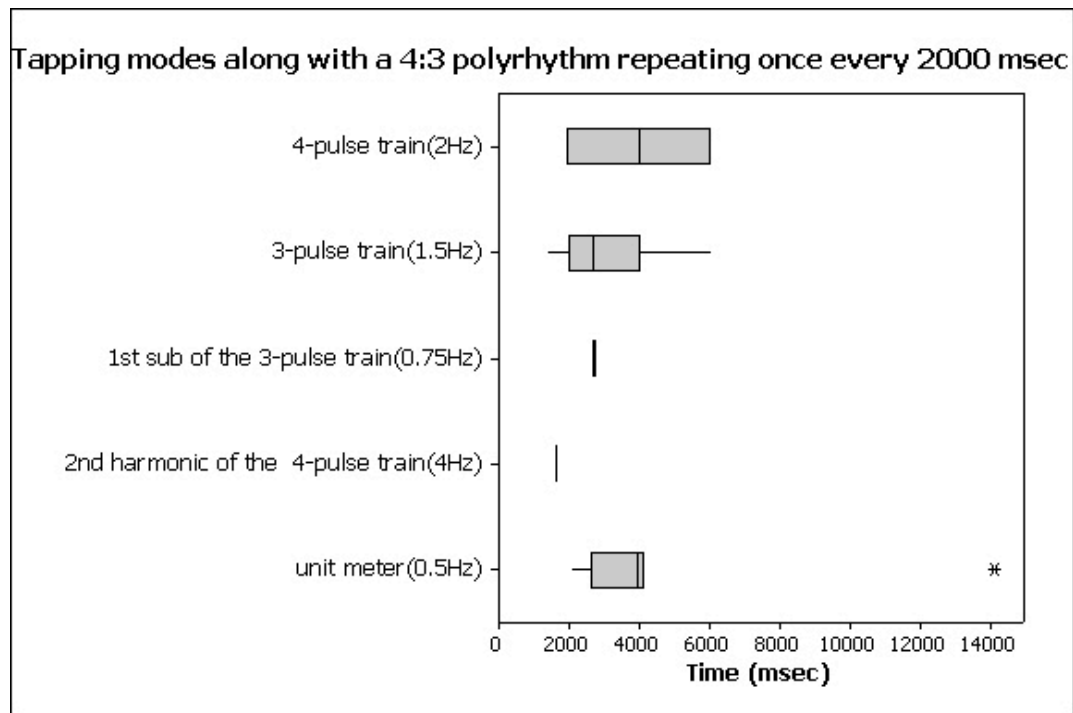


Figure 8.9: Multiple boxplots for the range of tapping modes and their corresponding period of time within which people start tapping along with a 4:3 polyrhythm repeating once every 2000 msec.

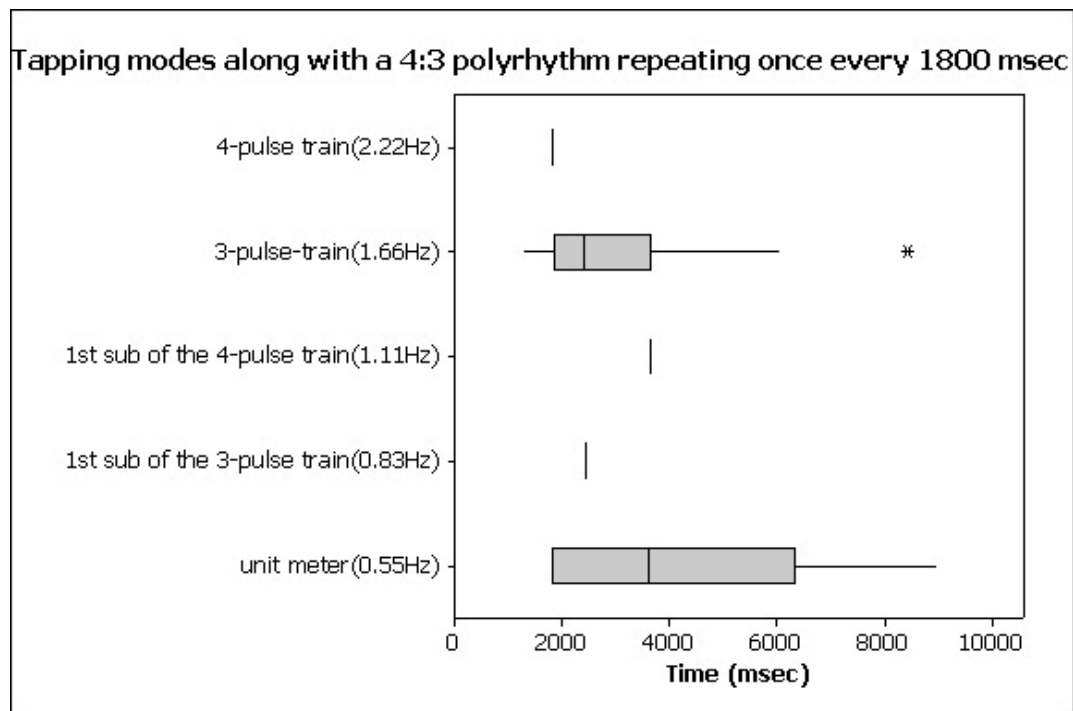


Figure 8.10: Multiple boxplots for the range of tapping modes and their corresponding period of time within which people start tapping along with a 4:3 polyrhythm repeating once every 1800 msec.

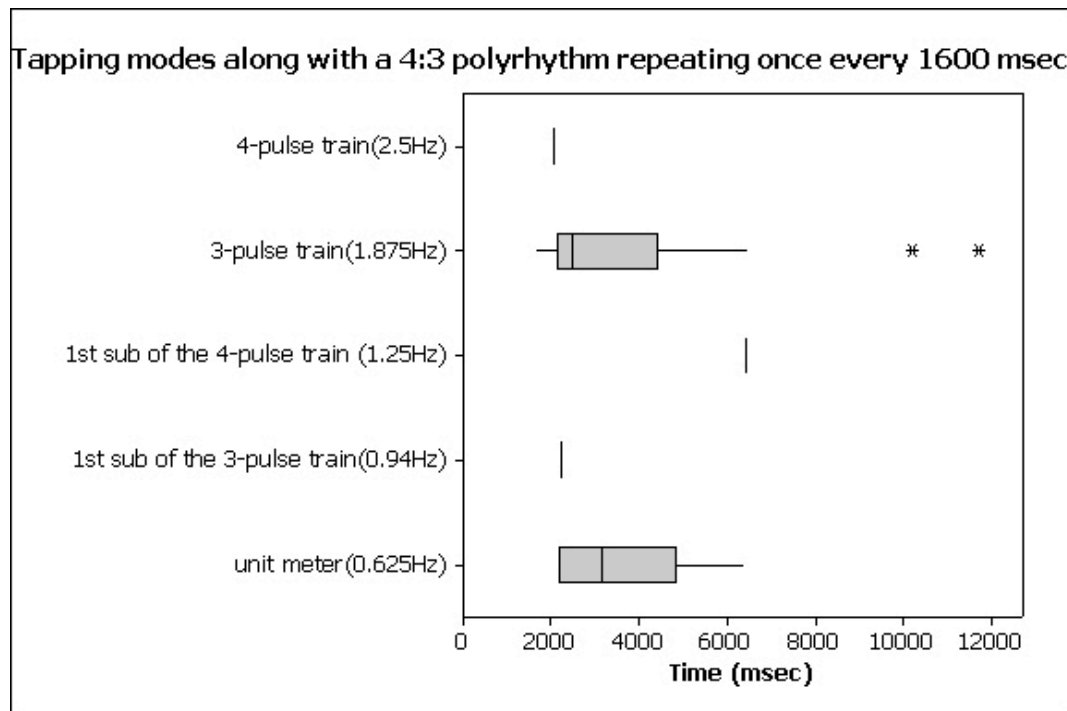


Figure 8.11: Multiple boxplots for the range of tapping modes and their corresponding period of time within which people start tapping along with a 4:3 polyrhythm repeating once every 1600 msec.

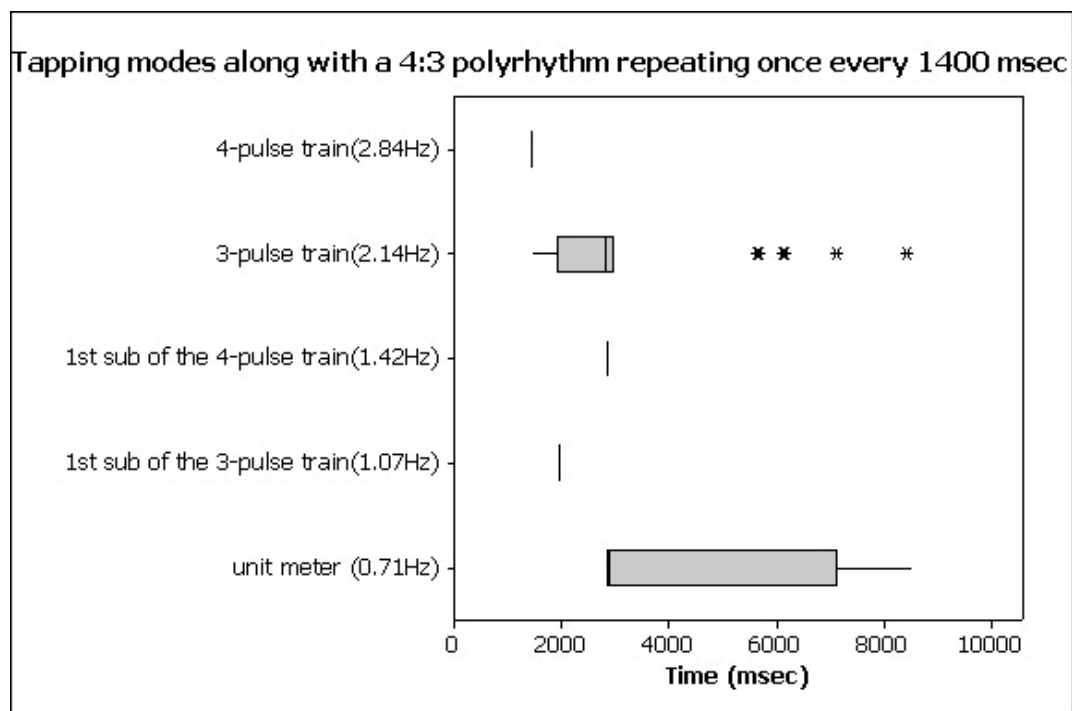


Figure 8.12: Multiple boxplots for the range of tapping modes and their corresponding period of time within which people start tapping along with a 4:3 polyrhythm repeating once every 1400 msec.



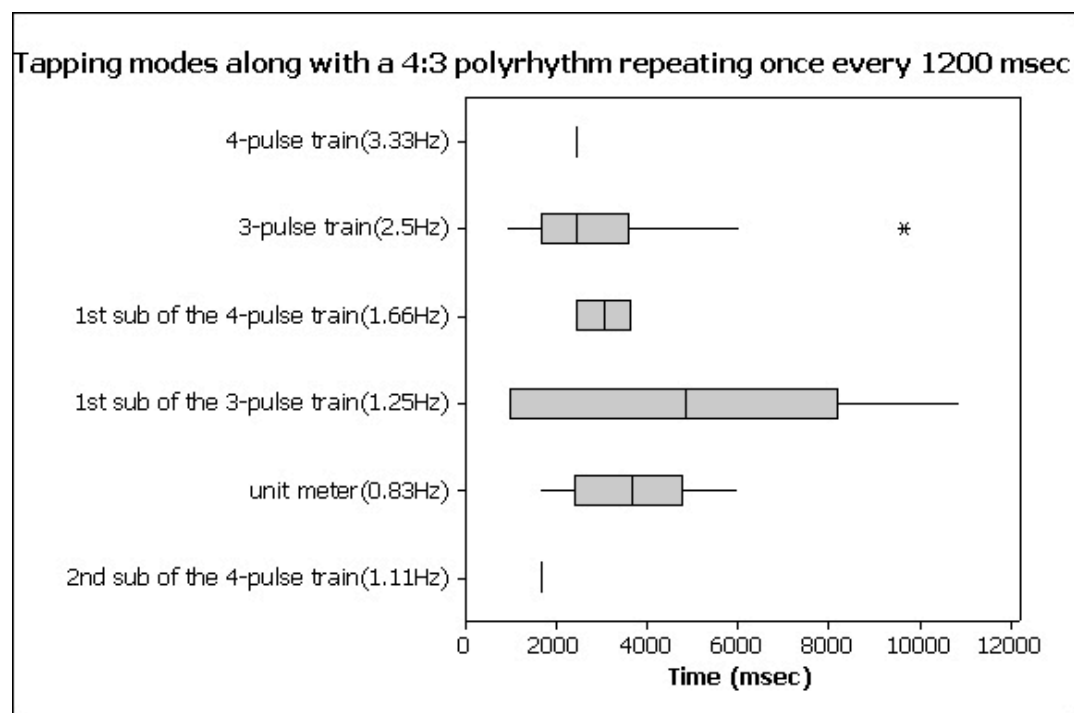


Figure 8.13: Multiple boxplots for the range of tapping modes and their corresponding period of time within which people start tapping along with a 4:3 polyrhythm repeating once every 1200 msec.

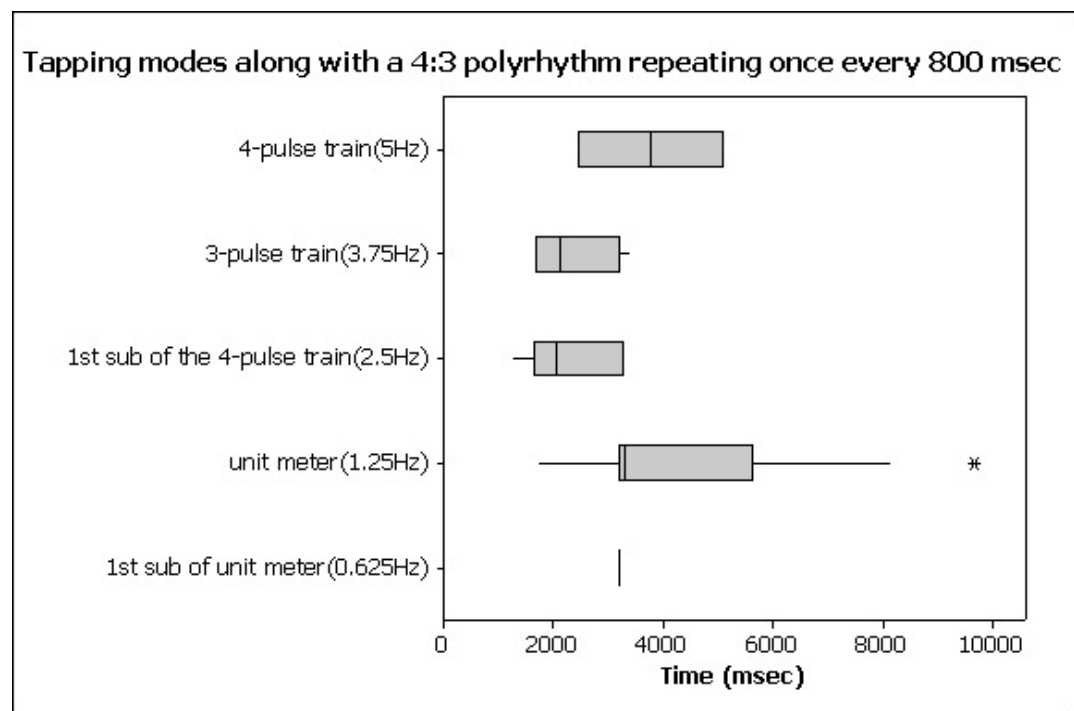


Figure 8.14: Multiple boxplots for the range of tapping modes and their corresponding period of time within which people start tapping along with a 4:3 polyrhythm repeating once every 800 msec.

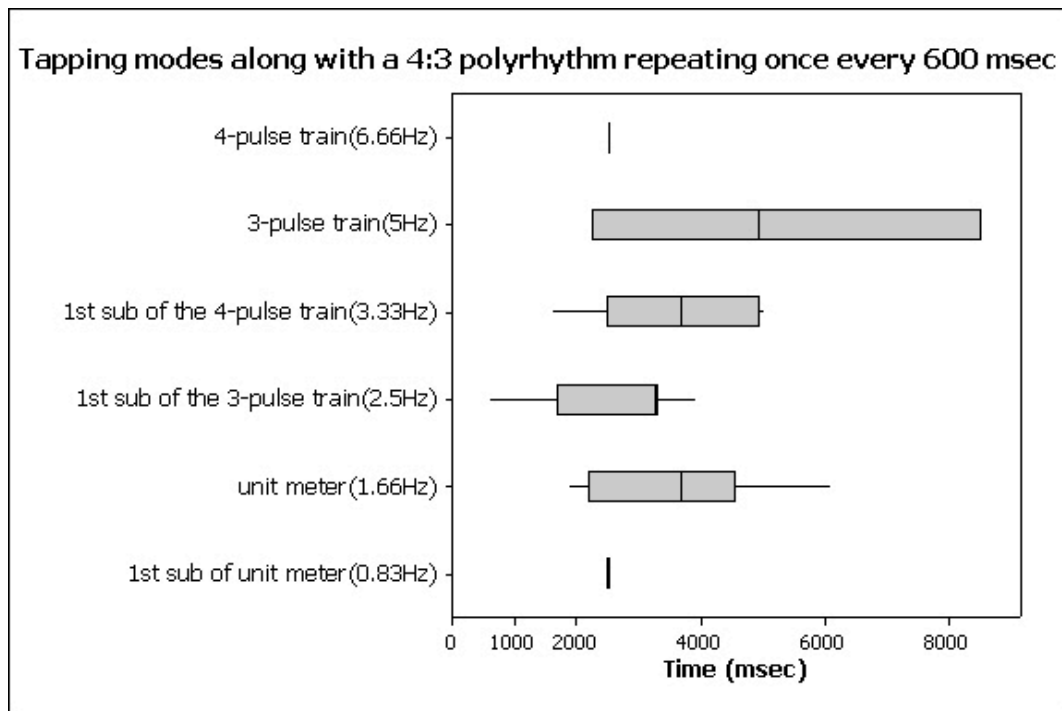


Figure 8.15: Multiple boxplots for the range of tapping modes and their corresponding period of time within which people start tapping along with a 4:3 polyrhythm repeating once every 600 msec.

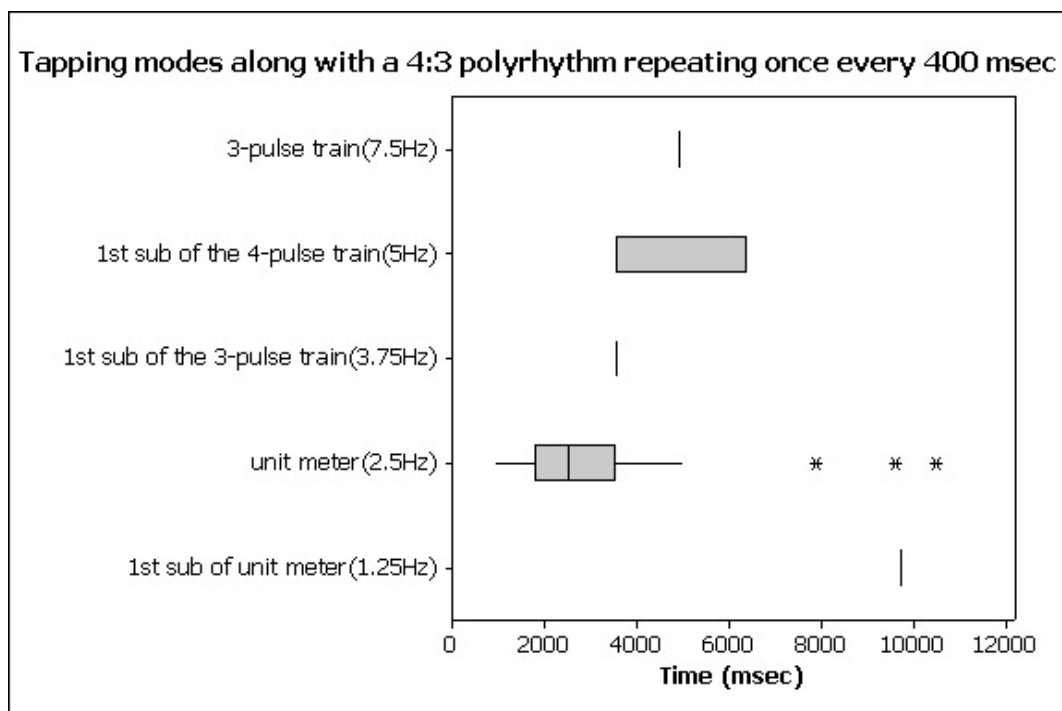


Figure 8.16: Multiple boxplots for the range of tapping modes and their corresponding period of time within which people start tapping along with a 4:3 polyrhythm repeating once every 400 msec.

#### 8.4 EVALUATION OF THE MODEL'S PREDICTIONS

The third and last section of this chapter focuses on examining whether the predictions on human tapping behaviour made by the neural resonance model (described in Section 7.1) can be observed in the empirical data presented in the previous section. The comparison is two-fold, one concerning the categories of tapping and the other the period of time it takes for people to start tapping. The section begins by comparing the predicted categories of tapping with those found in the experiments.

##### 8.4.1 *Extension of categories of tapping behaviours*

A quick review of the boxplots presented in the previous section reveals that in general, tapping behaviours agree with the findings reported by Handel and Oshinsky (1981). In other words, people tap along with either of the pulse trains, along the co-occurrence of the two pulse trains, and along with every other element of the pulse trains. However, there are cases where the tapping behaviour performed by people goes beyond these categories, but yet is in line with the predictions suggested by the neural resonance model. In particular, these are categories of tapping behaviours that correspond to: tapping the 2<sup>nd</sup> harmonic of the 4-pulse train when the pattern is repeated once every 2500 msec (5 observations) and 2000 msec (1 observation).

Also, there are additional empirical observations of tapping behaviours not observed in the Handel and Oshinsky, which are less frequent at least with regard to the rest of the observations, but nonetheless they imply a pulse like tapping along with the presented polyrhythm. Furthermore, for these categories of tapping behaviours there are no frequency responses

in the neural resonance model that could be associated with. In these behaviours people choose to tap the first subharmonic of the unit pattern when the stimulus is repeated once every 800 msec (1 observation), 600 msec (2 observations) and 400 msec (1 observation). Another novel empirical tapping behaviour can be found in the case where the 4:3 polyrhythm is repeated once per 1000 msec, where the participant groups 6 pulses of the 4-pulse train. The inter-onset interval of those pulses is  $1000:4 = 250$  msec, therefore the participant produces a pulse like response with a period of 1500 msec, or expressed in terms of rhythmic frequency 0.666 Hz. The second observation is related to the 4:3 polyrhythmic pattern repeating once every 1200 msec. The participant groups three pulses of the 4-pulse train. The inter-onset interval of those pulses is  $1200:4=300$  msec, therefore the participant produces a pulse like tapping with a period of 900 msec, or 1.11 Hz. This tapping behaviour is similar to the one described above in a sense that they both group pulses of the 4-pulse train. In the latter case 3 pulses are grouped while in the former 6.

In general, on the one hand, the model correctly predicts the extension of the categories of tapping behaviours to include for example 2nd harmonics of a pulse train. On the other hand, the experiment has identified new but relatively rare ways of how people choose to tap periodically along with a given polyrhythm and for these cases there is no prediction from the model associated with them.

#### 8.4.2 *Evaluating oscillator relaxation time*

The second part of the evaluation is concerned with the neural resonance model's predictions about the relative time it takes for people to start tapping in different modes. The rationale of the comparison is a

modification of a hypothesis suggested by Large (2000) according to which the relaxation time of an oscillator can be used to simulate the time it takes for people to start tapping - as soon as they feel a beat - along with a given rhythm. Large (2000) tested his hypothesis by comparing with his model the time it took for humans to start tapping along with a ragtime stimuli. For this purpose, the relaxation time of a network of oscillators was compared with beats to start tapping (BST) (Snyder and Krumhansl, 2000) explained in detail below. Based on the outcome of this comparison Large (2000, p. 555) argues "the model's relaxation time squares reasonably well with human performance". In this section, a similar comparison is undertaken for the case in which people are tapping to polyrhythms.

A note should be made to consider the difference between the time needed for people to perceive a beat and the time it takes to translate the perceived beat into an action. Firstly, we may reasonably assume that beat perception precedes tapping to the beat. Secondly, since the neural resonance model simulates the development of neural dynamics that give rise to the perception of beat, then the rationale suggested by Large (2000) whereby tap time is directly compared with oscillator dynamics may need to be adjusted to accommodate the time difference between perception and action. In the experiments described in the previous section a modest effort has been made to minimise the time difference between perception and action by emphasising to the participants to start tapping as soon as they feel the beat.

Before we continue with the evaluation of the prediction with which this section is concerned, we juxtapose the similarities and differences between the set up of the study (e.g., model details and experimental variables) adopted in this thesis with the one in the Large (2000) study. Furthermore, we give a detailed version of the argument that a comparison between

the time it takes for people to start tapping with the time it takes for an oscillator to relax as conducted by Large (2000) should be preferably conducted by focusing on a timescale rather than exact points in time.

#### 8.4.2.1 *Differences and similarities between studies*

The network of oscillators in the Large (2000) study is similar to the neural resonance model used in this thesis. It is important to note that the relaxation time of the oscillators depends on the numerical values given in the free parameters of the model (Large, 2000, p. 541-545), however Large states that they were not explicitly adjusted to maximize goodness of fit to the empirical data described below. The relaxation times of the simulations presented in this section are obtained using the nominal values described in Chapter 7.

There is a slight difference between the two set-ups in the way the time it takes for people to start tapping is quantified. In Large's study the source of empirical data is from the Snyder and Krumhansl (2000) study. There the time it takes for people to start tapping is expressed in terms of beats to start tapping. In order to reflect this, Large expresses the time it takes for an oscillator to reach relaxation in terms of number of beats. So for example if one of the resonating oscillators has a period of 250 msec, and 1500 msec are needed for the oscillator to relax, then the latter divided by the former provides the time is needed to relax in terms of beats. So for this particular oscillator  $1500 : 250 = 6$  beats are needed before the oscillator reaches relaxation. In this thesis however, neither the time it takes for people to start tapping nor the time it takes for oscillators to relax is expressed in terms of beats but instead it is expressed in time units (msec).

#### 8.4.2.2 *Limitations in the method of testing the original hypothesis*

In general, there is no reason to doubt that a modelled oscillator reaches relaxation at some fixed point in time (given a particular set of initial conditions). Large's results in testing his hypothesis support that the above time squares well with the mean time it takes for people to start tapping. However, in the following three paragraphs we identify weaknesses and limits in Large's arguments and claims.

As previously noted, Large (2000) refers to the Snyder and Krumhansl (2000) study in order to obtain empirical data regarding the time it takes for people to start tapping. In that empirical study there are three groups of participants reported. Two of the groups consist of musically experienced participants and one group of non-experienced. All three groups are asked to identify and tap a beat along with a ragtime musical stimulus. In the first group of 8 musically experienced participants Snyder and Krumhansl report a mean BST time of 5 and 6 beats for pitch varied and monotonic versions of the stimulus respectively. In the second group, the 12 musically experienced participants were asked to perform the same task while listening to the same stimulus. The mean BST time was 3 to 4 beats in this case. Snyder and Krumhansl report a very low standard error for the both mean BST times, meaning that this sample mean time is quite close to the population mean. Lastly, the mean BST time for the group of 8 musically inexperienced participants was 9 and 11 beats for pitched-varied and monotonic versions of the ragtime stimulus respectively. Taken together for both experienced and inexperienced participants the mean BST ranges from 3 beats to 9 beats for the pitch-varied version of the ragtime stimulus, and 4 to 11 for the monotonic version. Despite this variation on the mean BST time, Large focuses on the mean BST of the particular group of 12 experienced participants to conduct the comparison

with the BST time predicted by the model, and thus, his argument that the model's relaxation time squares reasonably well with human performance relies on choosing a particular group of participants.

Furthermore, Large (2000) focuses exclusively on the mean BST time and compares it with the exact point in time at which an oscillator relaxes. However, a mean BST time is a measure of location that summarises the variation of BST time observed among participants. In fact, Snyder and Krumhansl report variation with regard to the BST time. For example, for the specific experiment, which is also the reference point for the Large (2000) study, Snyder and Krumhansl report that 75% of the musically experienced subjects (9 out of the 12) have started tapping after 4 beats, while 10 beats are needed for 95% of the subjects to start tapping. From that perspective a measure that focuses on the variation of BST time would be more appropriate than the mean BST time. Also, we have already mentioned that the empirical data regarding the absolute time to start tapping do not conform with a normal distribution, thus calculation of mean may be inappropriate. A quick way to check that our own empirical data are not distributed normally is by observing that the boxplots are not symmetrical in general even for those cases where a relatively large number of observations are associated with a particular boxplot. For example, in Figure 8.6 (repeated below as Figure 8.17) the boxplot corresponding to people start tapping along with every other element of the 4-pulse train (1st subharmonic of the 4-pulse train) contains 15 observations and it is not symmetrical. Another example, where the frequency of a particular tapping behaviour exceeds 25 observations (Figure 8.9 - Boxplot for 3-pulse train - 26 observations) the boxplot for that particular behaviour implies a right skewed distribution, thus the use of a mean is certainly not the best numerical summary for the center of that skewed distribution. In any case,



a larger sample of observations may be necessary to unveil the exact shape of distribution of the time it takes for people to start tapping along with a given rhythm in general, thus to inform which type of summary statistics should be used.

For these reasons, we suggest that trying to compare the mean time it takes for people to start tapping with the particular time at which an oscillator reaches relaxation may not be the most appropriate way to support the hypothesis whereby a model like the neural resonance model can make predictions about timing aspects of human tapping behaviour. Instead, we suggest that the focus should be on the temporal development of the oscillator dynamics, which as we will see, shares the same timescale with the period of time within which people start tapping to rhythms. In this way, the neural resonance model can be used to make plausible predictions about timing aspects of human tapping behaviour.

#### 8.4.2.3 *Evaluation*

Figure 8.1, which is repeated below as Figure 8.17 for convenience, shows the development of the oscillatory dynamics over time in the presence of a 4:3 polyrhythm with 1 Hz repetition frequency. For some oscillators the transient stage (the stage until a relaxation is reached) is longer than others. For example, the oscillators with frequencies 3 Hz and 4 Hz relax after 4 seconds, while the oscillators of 2 Hz and 1 Hz frequency are still in their transient stage in the 4th sec of the stimulation. Also, the oscillators that reflect the metrical structure of the 4:3 polyrhythm (e.g., 4Hz, 3Hz, 1Hz, etc.) start to resonate from the very early stages (after the 1st second) of the stimulation and they stay in transience for some time before they reach relaxation.

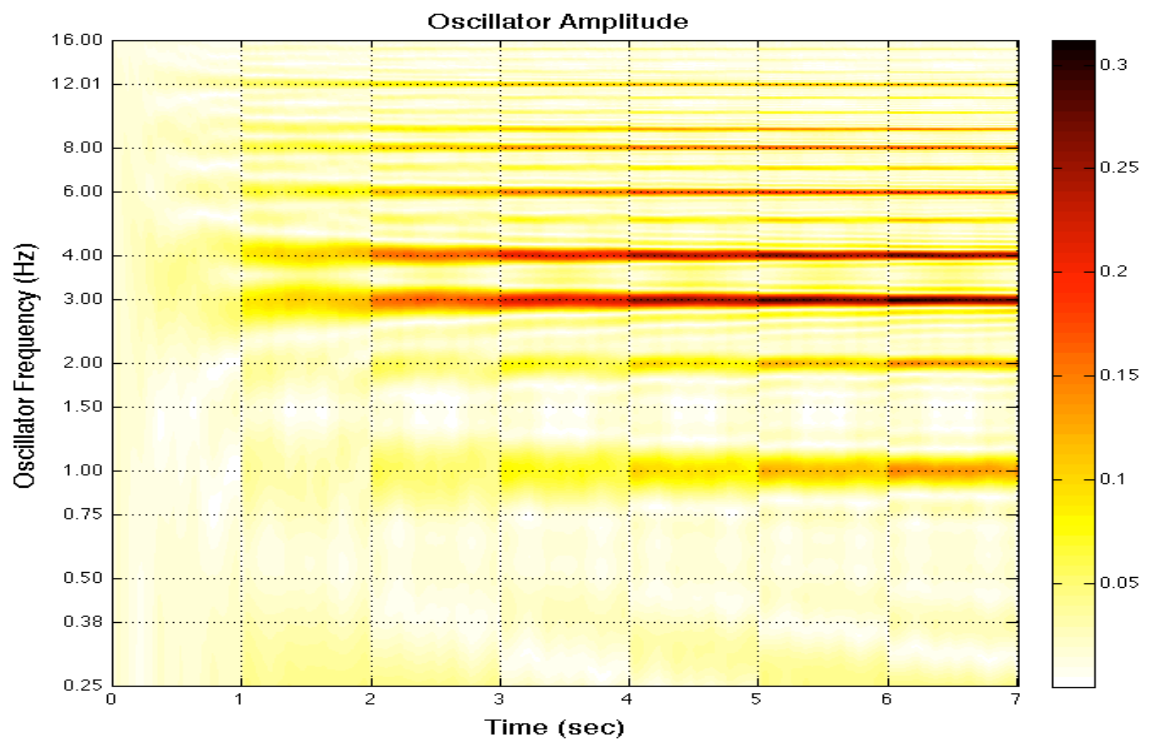


Figure 8.17: Temporal development of resonances among the bank of oscillators in the presence of a 4:3 polyrhythm repeating once per second. The period of time at which oscillators start to resonate begins almost immediately after stimulation. The transient stage (the time until relaxation is reached) is different for each oscillator, but nonetheless it spans a period of several seconds. For example, oscillators with frequencies around 3 Hz and 4 Hz relax within the first 4 seconds of the stimulation.

Figure 8.6 is repeated below to illustrate the periods of time within which tapping modes along with a 4:3 polyrhythm repeating once per second take place. There is no distinction between experienced and inexperienced participants. We can see that the periods of time to start tapping share the same timescale to the periods of transient dynamics showed above. For example, for all 5 categories of tapping associated with this particular combination, people start tapping after 1000 msec. This is similar to the time it takes for the resonances to start developing as shown in the oscillators dynamics of Figure 8.1. Also, the temporal period within which each empirical tapping mode is observed is similar to the transient period in the oscillatory dynamics, i.e., first seconds after the stimulus is heard/presented.

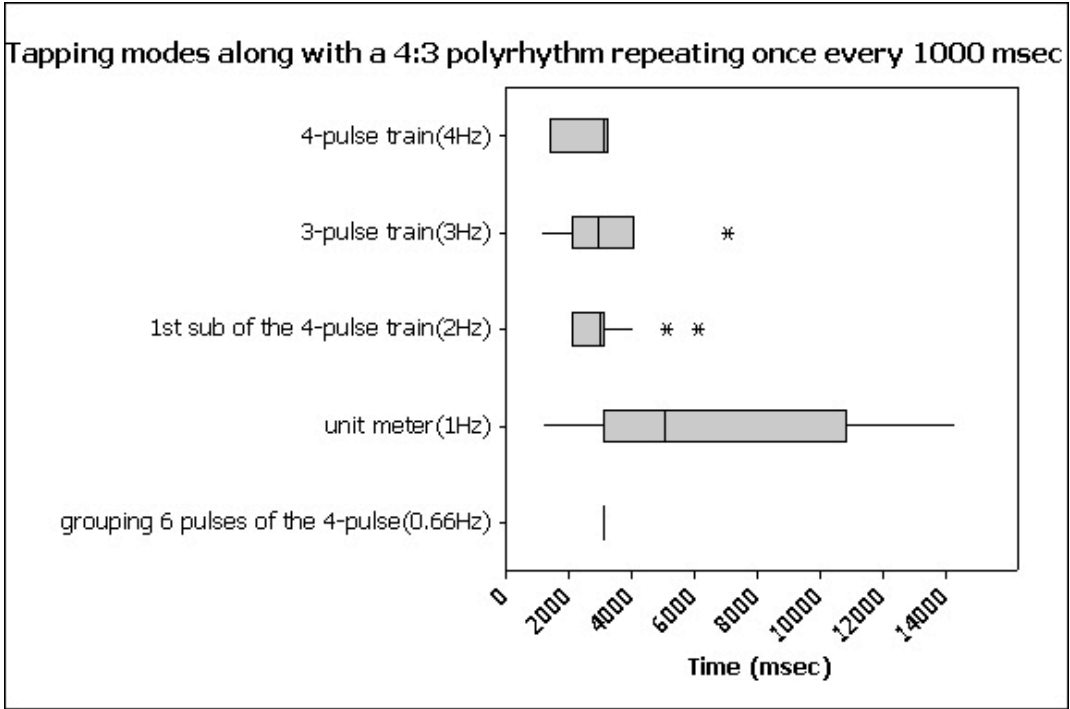


Figure 8.18: The period of time it takes for people to start tapping in a given mode along with a 4:3 polyrhythm repeating once every 1000 msec.

In general, the timescale of the period within which people start tapping along with a polyrhythm and the period of transient dynamics is of the

same magnitude for all types of polyrhythms and presentation rates. We further present empirical data related to two more types of polyrhythm, i.e., 3:2 and 2:5 repeating once every second, and the corresponding predictions by the model. An exhaustive presentation of graphs covering the 50 combinations is achievable, however it is not included here for purposes of concise presentation.

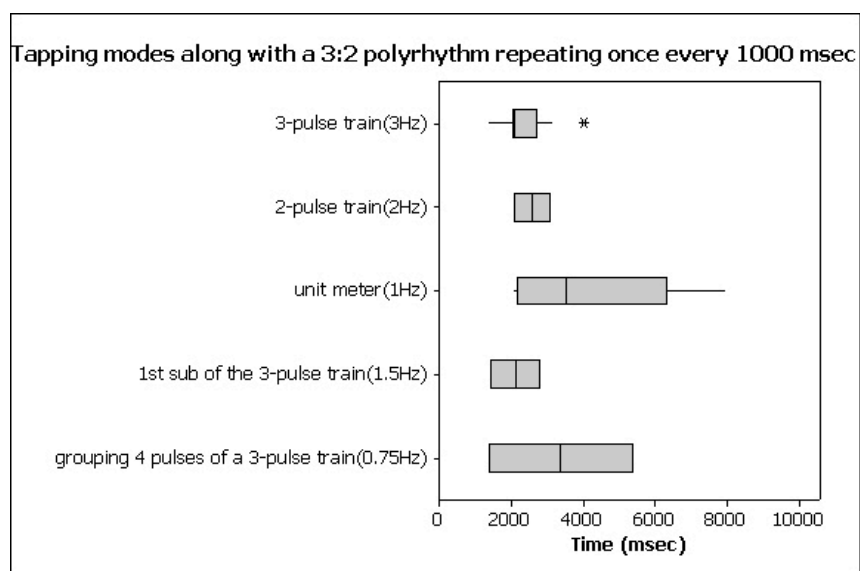


Figure 8.19: The modes of tapping and their corresponding periods of time (boxplots) within which people start tapping along with a 3:2 polyrhythm repeating once every 1000 msec.

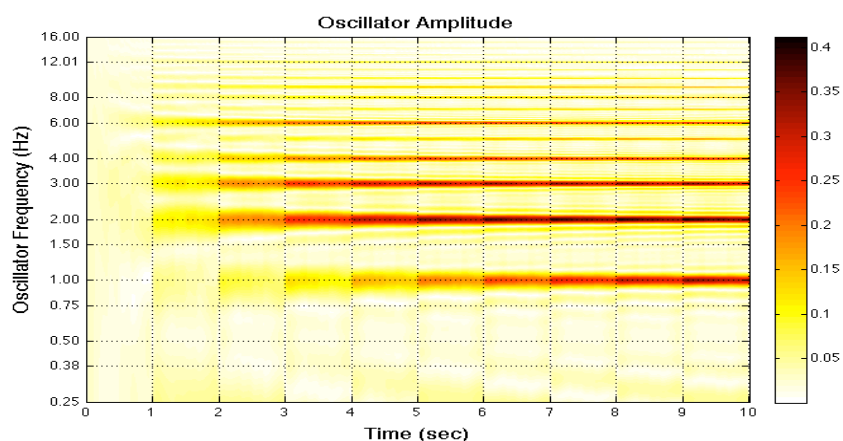


Figure 8.20: Temporal development of resonances in the network in the presence of a 3:2 polyrhythm repeating once every 1000 msec.

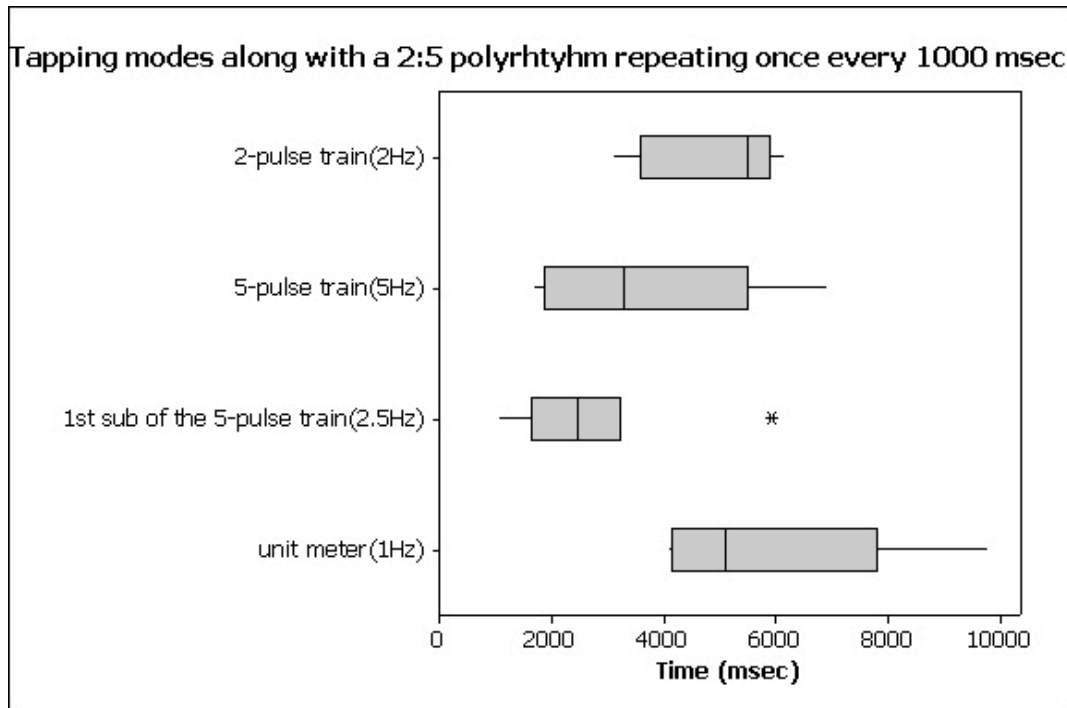


Figure 8.21: The modes of tapping and their corresponding periods of time (boxplots) within which people start tapping along with a 2:5 polyrhythm repeating once every 1000 msec.

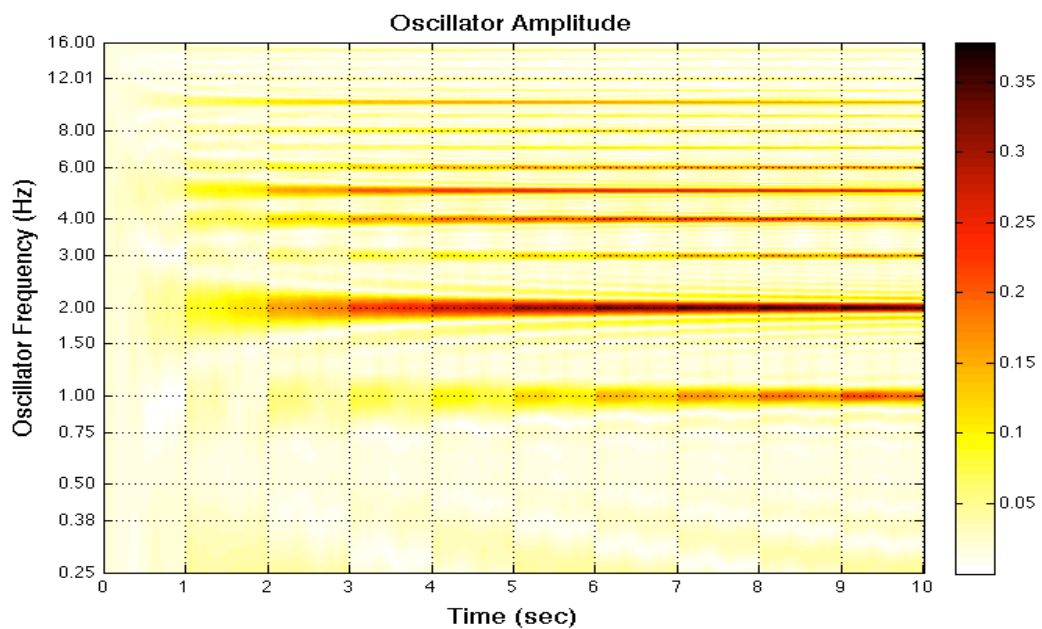


Figure 8.22: Temporal development of resonances in the network in the presence of a 5:2 polyrhythm repeating once every 1000 msec.

## 8.5 SUMMARY

In this chapter we used our previous analysis of the response of the neural resonance model to polyrhythms to create two new classes of prediction about human behavior in tapping to polyrhythms. In order to test these new predictions, we carried out a new empirical human study on tapping to polyrhythms. This new study collected substantially more data and analysed it in more depth than the previous existing study due to Handel and Oshinsky (1981).

The first of the predictions posited that the general categories of tapping behaviours along with a polyrhythm reported by the study of Handel and Oshinsky, was incomplete, and that specific new kinds of human tapping behavior should be observable. These predicted and previously undocumented human tapping behaviours were then subsequently observed. In the second prediction we referred to a hypothesis due to Large (2000) whereby a connection can be established between the evolution of oscillator dynamics and time taken for people to start tapping along with a rhythm as soon as they feel a beat. The need for a modified version of Large's hypothesis was argued and a modified, more defensible version was framed accordingly. In particular, this more plausible version of the original hypothesis suggested that the timescale of a relaxation, as opposed to an absolute point in time where relaxation occurs, is more appropriate for measuring the phenomenon. The modified version of the hypothesis suggested a novel set of summary statistics for evaluating the revised hypothesis.

To reflect on this summary in a little more detail the nature of predictions of the model led to conducting a series of novel empirical experiments in polyrhythm perception in order to acquire the necessary data. In particular,

we have conducted an extended version of the study reported by Handel and Oshinsky (1981) with many more data points, supporting a deeper level of analysis. A range of tapping modes were identified and the time it took for people to start tapping along with the polyrhythm was analysed. Subsequent comparisons between empirical data and predictions showed that the model correctly predicted the extension of the categories of tapping to include 2nd harmonics of the input frequencies and subharmonics of the frequency repetition of the pattern, and also that the period of the temporal development of the dynamics in the model is relative to the period of time within which people start tapping along with a polyrhythm.

Lastly, the empirical data have shown tapping modes that are not predicted by the neural resonance model nor they have been found in previous empirical studies, a finding that complements further our understanding of the way people perceive polyrhythms. A series of limitations have been exposed and interpreted as opportunities for future work.





## CONCLUSIONS, LIMITATIONS AND FURTHER WORK

---

This thesis has tested and analysed the theory of neural resonance, a physiologically plausible theory of human rhythm perception, in the context of modeling certain aspects of human polyrhythm perception. According to the theory, human rhythm perception is directly related to neurodynamics (Large, 2008). We have investigated the extent to which the theory accounts for certain aspects of human polyrhythm perception, in particular the way people tap along with polyrhythmic stimuli in a periodic way and the time it takes for them to start tapping. The former human behaviour is associated with the identification of a beat in a polyrhythmic stimulus and the latter is associated with the time it takes to perceive a beat. We have used the theory to make new predictions about human behaviour and we have carried out new empirical studies to test those predictions. In the following three sections we review the results, discuss the limitations and identify directions for future work.

### 9.1 CONCLUSIONS

In general, we have focused on two behavioural aspects of human polyrhythm perception which the neural resonance model addresses. In particular, the neural resonance model responds to appropriate stimuli with responses that reflect both the different ways in which people tap regularly along with polyrhythmic stimuli, and the timescale according to which they

begin to tap. The methodology of this work involves stimulation of the model with different polyrhythmic stimuli at different tempi, and analysis both of the frequency response pattern and of the development of the resonant dynamics. In particular, the proposed methodology for studying the second aspect of the human behaviour described above, constitutes an original alteration to an existing methodology for concrete reasons that have been described in Chapter 8. Briefly, we proposed that the timescale of relaxation should be considered rather than a specific point in time where relaxation occurs, and secondly, we proposed a different set of summary statistics - to describe empirical data - as more appropriate to those utilised by the existing methodology. Additionally, we checked for consistency in the model's behaviour by systematically manipulating the numerical values of the parameters of the state variable, the initial conditions of the oscillators, and the amplitude of the polyrhythmic stimulus, and found consistency of results for a range of numerical values (within an appropriately defined range). This methodology has originally been developed for the purposes of this thesis and to our best knowledge it is the first time that such an analysis is published. Furthermore, we examined the degree to which the neural resonance model behaves similarly to the less abstracted and more directly biologically rooted model of Wilson and Cowan (1972), in order to limit the degree of uncertainty of the former regarding its biological assumptions. We have found that the two models behave similarly and thus, concluded that we can place a reasonable limit on uncertainty with regard to the biological assumptions made by the neural resonance model. The theory of neural resonance unlike many theories relates high level perception, like rhythm perception, to behaviour of neurons, so it is important to test it in order to explore its limits and strengths in our effort to understand the

human brain. As imaging techniques become widely spread in the future new techniques will become available to test the model. Nonetheless the existing research has unearthed previously unseen human behaviour thus making the theory of neural resonance a good candidate for examining potential physiological plausible ways of how high level behaviour can be explained in terms of brain dynamics. In the introduction of this thesis we have underlined the importance of rhythmic stimuli in regulating brain activity and briefly discussed the therapeutic applications of interventions that utilise rhythmical stimuli to treat various psychological and biological diseases associated with brain malfunction. Theories of rhythm such as the neural resonance theory may lead to a better understanding of the physiological underpinnings of rhythm perception which could further lead to the development of carefully tuned and customised interventions to promote human well-being in a non-invasive way.

#### 9.1.1 *Validation against existing empirical data*

The first research question seeks to examine whether the neural resonance model can account for existing empirical data in polyrhythm perception. The question as stated in Chapter 1 is repeated below for convenience.

Question 1: To what extent can the theory of neural resonance provide an account for existing empirical data on how humans perceive polyrhythms?

In order to address the above question we conducted a series of tests with the model. The design of the tests mirrored existing experimental studies in human polyrhythm perception, and took advantage of existing empirical data about how people perceive the beat in polyrhythmic stimuli

under different circumstances. This allowed the behaviour of the model to be compared with human behaviour. In particular, our tests mirrored the experimental study of Handel and Oshinsky (1981) in which people were asked to tap along with a polyrhythm in a regular pulse-like way. Handel and Oshinsky presented five different types of polyrhythms to people at ten different tempi. We used the same polyrhythmic stimuli to stimulate the neural resonance model. Our tests showed that the neural resonance model produced frequency response patterns whose categories matched the full range of human tapping behaviour reported by Handel and Oshinsky. This was the case for diverse polyrhythms and diverse tempi as used by Handel and Oshinsky .

#### 9.1.2 *Evaluation of model's predictions*

The second research question seeks to examine the model's ability in predicting aspects of human behaviour in polyrhythm perception. The question as originally posed in Chapter 1 is repeated below.

Question 2: Can we make novel predictions about human behaviour in polyrhythm perception based on the theory of neural resonance and subsequently test these predictions empirically?

We saw in Chapter 6 that when the model is stimulated with polyrhythm parts of the generalised frequency response pattern includes resonances not identified in Handel and Oshinsky's studies of human behaviour. Some of these resonances correspond to tempi faster than humans can tap, according to limits discovered in sensorimotor synchronization studies (see Chapter 5, Subsection 5.2.2). However, other resonances fall in the humanly

tappable range. These resonances can be thought of as predictions made by the model about human behaviour. More specifically, the predictions suggest that (subject to tempo limits) human tapping behaviour should include tapping responses that reflect 2<sup>nd</sup> harmonics of the fundamental frequencies of the polyrhythm. We carried out a larger scale, more detailed version of the experiments conducted by Handel and Oshinsky, and found humans behaving in the way predicted by the model and overlooked by the earlier experimenters. More specifically, we identified tapping frequencies that correspond to 2<sup>nd</sup> harmonics of a 4-pulse train of a 4:3 polyrhythm repeating once every 2500 msec and 2000 msec. We also framed and tested a second prediction about human tapping behaviour concerning the relative time it takes for people to start tapping along with different polyrhythms (Chapter 8). This prediction was derived from the relative time it takes for an oscillator to reach relaxation in the presence of different rhythmic stimuli. This prediction was also tested by conducting novel experiments with people tapping to polyrhythms. More specifically, the relative time it took for people to start tapping along with a polyrhythmic stimulus was monitored for all five types of polyrhythms and for all repetition rates reported by Handel and Oshinsky. To the best of our knowledge, our experiment appears to be the first occasion in any literature that the relative time to start tapping along with a polyrhythm has been empirically tested. Our analysis showed that the relative timescale within which people start to tap along with different polyrhythms matches the relative timescale within which the resonances of the oscillator develop until they reach relaxation.

## 9.2 LIMITATIONS

In this section we discuss the limitations of the research.

1. The model predicts *categories* of tapping but fails to predict the *relative popularity* of those categories at different tempi (Chapter 6). A reasonable candidate hypothesis (based on some neurophysiological evidence - subsection 6.2.3) would be to use the magnitude of the amplitude response of the different oscillators in the model to predict the relative preferences of the listeners for different beats in a polyrhythm. In other words, the strongest amplitude in the frequency response pattern of the model might be expected to predict the most popular human preference for a particular way of tapping, and so on for the next strongest amplitude, etc. In fact, such a prediction fails to match the human *relative popularity* of tapping categories.
2. The model not only fails to track the *relative popularity* of tapping categories at different tempi - it features a limitation that appears to virtually guarantee this failure. The problem is that the relative popularity of human tapping categories *changes* for different tempi. By contrast, in the case of the model, the effect of different tempi is simply to shift the amplitude pattern of the frequency response along the x-axis (frequency axis). In other words, when humans tap to a polyrhythm in a regular way, the preferences of tapping observed are a function of *absolute tempo*. For example, for fast presentation rates of n:m polyrhythms there is a higher preference in tapping along with the co-occurrence of the two constituent pulse trains than for slow presentation rates (Handel and Oshinsky, 1981). However, when stimulating the model with an n:m polyrhythm (Chapter 6) at various tempi the frequency response pattern of the model remains invariant in terms of the relative amplitude levels.

3. The model does not predict all human categories of tapping, albeit it represents all common ones. More specifically, in our own experimental studies reported in Chapter 8 we identified some very rare tapping behaviours in which participants grouped six pulses of the 4-pulse train of a 4:3 polyrhythm repeating once per 1000 msec, and grouped three pulses of the 4-pulse train of a 4:3 polyrhythm repeating once per 1200 msec. These behaviours, although rare, are not predicted by the model.
4. In this research, we chose to configure the neural resonance model using a single bank of oscillators and without interconnectivity among oscillators. There is some evidence that the perception of rhythm takes place in multiple, spatially distinct brain regions (Chen et al., 2008; Karabanov et al., 2009). Thus there can be arguments for more complex models, using interconnected multiple banks of oscillators. However, the model we have chosen has the advantage of simplicity and parsimony. Also, it is unclear how more complex models should be parameterized within a very large configuration space while retaining physiological plausibility.
5. Although the model accepts input in a variety of forms (Chapter 5 and 7), in each case, physical characteristics of the stimulus such as pitch and timbre have been abstracted from the stimulus, leaving bare rhythmic information. Consequently, the model gives little indication how these musical dimensions might affect rhythmic processing. (Jones, 1987).

### 9.3 FURTHER WORK

1. The existing model predicts categories of tapping but the amplitudes of resonances do not predict the relative popularity of those categories at different tempi. It may be unreasonable to expect that they should, but if we make the assumption (based on some evidence provided by brain



imaging studies) that some hidden connection exists, then various issues need resolving. For example, the effect of tempo changes on the model is unlike the effect on human tapping behaviour, as the amplitude pattern of the frequency response is simply shifted along the x-axis. It could be that the model correctly models the state of oscillators in the brain, but that other factors moderate tapping behaviour such as musculo-skeletal capabilities of how fast one can tap, or it could be that our model of oscillators in the brain is wrong. One useful project would be to see if either of these hypotheses can be supported or ruled out by the application of brain imaging techniques.

2. One family of useful future research projects would be to explore the extent to which extended models using interconnected banks of oscillators might be able to model certain aspects of human rhythm perception more faithfully than the single bank model of oscillators used in the present work. For example, non-connectivity focuses the response of the network exclusively on external rhythmic stimulus (Large et al., 2010), whereas in some cases a better model might include influences from other processes (e.g. motor activity, or remembered rhythmic patterns). Such processes might be modelled by multiple connected banks of oscillators. There is some physiological justification for such variants of the model. Recent brain-imaging studies (Chen et al., 2008; Karabanov et al., 2009) suggest that rhythmic information is represented across broad cortical and subcortical networks and that both auditory and motor areas play key roles in both rhythm perception and production. Velasco (personal communication), suggests that employing more than one bank of oscillators might be used to model two communicating distant areas of the brain, such as the primary auditory cortex and supplementary motor area, with the output of one bank be employed to stimulate another. Clearly, such directed stimulation

could be used in principle to model afferent connections where peripheral information is carried inwards towards parts of the central nervous system, e.g., parts along the auditory pathway and the auditory cortex. Brain imaging techniques could be used to further constrain exploration of such models in physiologically plausible ways.

3. A related family of useful future research projects could explore the extent to which extended models using interconnectivity among oscillators in a single bank (as opposed to connections between banks of oscillators) might be able to model certain aspects of human rhythm perception more faithfully than the unconnected oscillators used in the present work. There is at least general physiological evidence for considering more complicated models. For example, interconnected oscillators within one network are a more accurate representation of the underlying connectedness of neural populations in the auditory nervous system than the non-connected oscillators are (Hoppensteadt and Izhikevich, 1996a). Tuning the model used in this thesis and comparing with human behavior might also help explore these issues. For example, as a preliminary step towards such a pilot, we briefly explored the frequency response pattern of the network in the presence of a 4:3 polyrhythm (Figure 9.1) with connectivity among oscillators in one network switched on, i.e., giving parameter (c) (Equation 5 - Chapter 4) an arbitrarily chosen small numerical value (0.1).

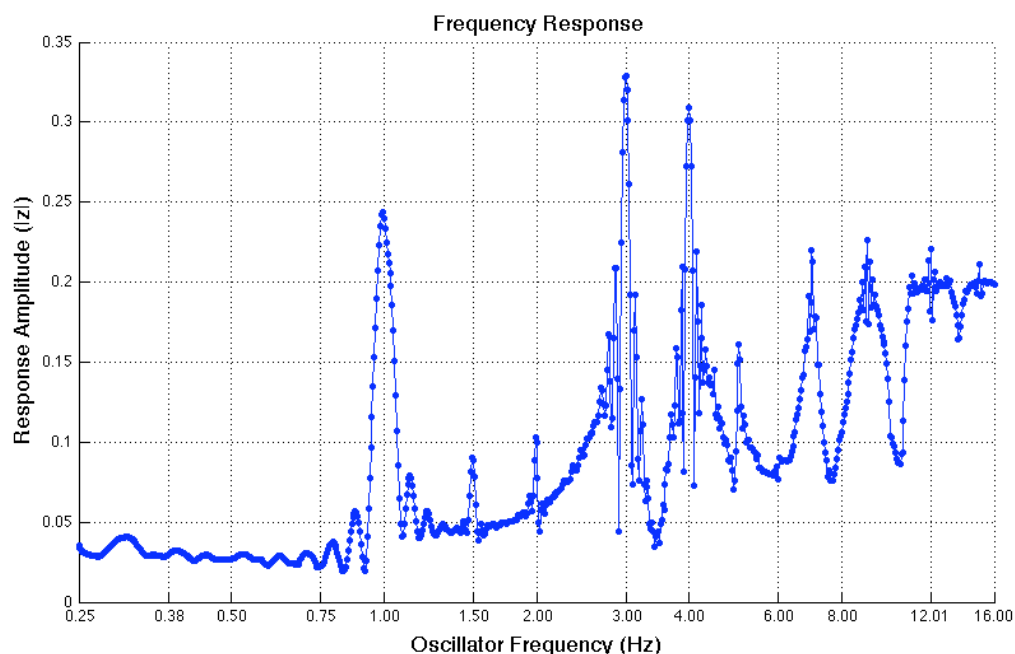


Figure 9.1: Frequency response of a network of interconnected oscillators in the presence of a 4:3 polyrhythmic stimulus of 1 Hz repetition frequency.

4. One way to model learning processes in rhythm cognition involves the use of connections between individual oscillators and between multiple oscillator banks. Loosely speaking, one could start with all oscillators weakly connected to each other, and to strengthen connections between oscillators whenever they oscillate at the same time, and to allow all other connection to atrophy. This version of Hebbian learning (Hoppensteadt and Izhikevich, 1996b) could model the way that learning appears to bias perception and action towards the familiar and away from the unfamiliar. Large (2011) has suggested that the neural resonance model can be adapted to demonstrate such processes, like for example the creation of tonal fields in the central nervous system that reflect different musical cultures. The design, implementation, testing and clear dissemination of new versions of the model for exploring these issues would constitute a valuable future contribution to knowledge.

5. A further possibility for future work lies in improving the way the

oscillators' relaxation time is estimated for the purposes of comparison with the time it takes for people to start tapping along with the polyrhythm. In the thesis, for pragmatic reasons, relaxation time is calculated by producing amplitude response graphs of the network in the presence of a polyrhythmic stimulus, and subsequently, by qualitatively observing the approximate point in time where the amplitude reaches its peak response and becomes stabilized. However, a more rigorous way to calculate the time to relaxation would be by expressing relaxation time as a function of time (Section 8.2). A small project taking advantage of this method would allow the identification of the exact point in time at which an oscillator relaxes, allowing more precise tests of the model.

6. The present thesis motivates various further empirical tests of human behavior. For example, we previously conducted empirical work (Chapter 8) to measure the time it takes to perceive a beat and tap along with polyrhythms. In particular we measured the absolute time it takes for people to start tapping along with polyrhythms. Snyder and Krumhansl (2000) made similar measurements using ragtime, but they focused on the number of beats passed before people started to tap. In specific cases, either measurement can trivially be transformed and expressed in terms of the other, but the general question of principle remains whether the time it takes for people to perceive beat is a function of absolute time or depends on a number of beats.

7. Empirical work shows that a number of physical characteristics of a stimulus are sources of interacting information from which people extract rhythmicity. For example, the preference of attending to a low-pitch pulse train in comparison to a high-pitch pulse train (of approximately similar tempi) when tapping, indicates that tonality information influences rhythm perception. Empirical studies using brain-imaging techniques can help us

understand how these empirically observed instances of rhythm perception translate into neural correlates. Results of such studies can help us refine our modeling aims to reflect an increased understanding of the exact neurophysiological process in a perceptual phenomenon like that of rhythm perception.

#### 9.4 FINAL REFLECTION

This thesis has exposed the theory of neural resonance to empirical and theoretical scrutiny by examining the extent to which it can be used to model aspects of human polyrhythm perception. We have shown that the neural resonance model can account for existing empirical data about how people perceive beat and meter in polyrhythms. We have used the model to make new predictions about human behavior. We have demonstrated empirically that humans behave in the predicted manner. We have analysed limitations of the model and theory, made comparison with competing models and theories, and identified physiologically motivated ways to refine the model. As a result, this thesis has cast new light on explanations of the phenomenon of rhythm perception directly in terms of brain dynamics, and identified likely paths to further illumination.

## BIBLIOGRAPHY

---

Victor Kofi Agawu. *African rhythm: a Northern Ewe perspective*. CUP Archive, 1995. (Cited on page 117.)

AA Andronov, EA Leontovich, II Gordon, and AG Maier. Theory of dynamical systems on a plane. *Israel Program of Scientific Translations, Jerusalem*, 1971. (Cited on page 50.)

D.G. Aronson, G.B. Ermentrout, and N. Kopell. Amplitude response of coupled oscillators. *Physica D: Nonlinear Phenomena*, 41(3):403–449, 1990. (Cited on page 48.)

Thaddeus L Bolton. Rhythm. *The American Journal of Psychology*, 6(2): 145–238, 1894. (Cited on page 20.)

Shannon Campbell and DeLiang Wang. Synchronization and desynchronization in a network of locally coupled wilson-cowan oscillators. *Neural Networks, IEEE Transactions on*, 7(3):541–554, 1996. (Cited on page 149.)

Ali Taylan Cemgil, Bert Kappen, Peter Desain, and Henkjan Honing. On tempo tracking: Tempogram representation and kalman filtering. *Journal of New Music Research*, 29(4):259–273, 2000. (Cited on page 39.)

Joyce L Chen, Virginia B Penhune, and Robert J Zatorre. Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex*, 18(12):2844–2854, 2008. (Cited on pages 199 and 200.)

- Russell M Church and Hilary A Broadbent. Alternative representations of time, number, and rate. *Cognition*, 37(1):55–81, 1990. (Cited on pages 28 and 34.)
- Eric F Clarke. Rhythm and timing in music. *The psychology of music*, 2: 473–500, 1999. (Cited on page 36.)
- Martin Clayton, Rebecca Sager, and Udo Will. In time with the music: the concept of entrainment and its significance for ethnomusicology. *European Meetings in Ethnomusicology*, 11:3–142, 2005. (Cited on page 20.)
- Charles E. Collyer, Hilary A. Broadbent, and Russell M. Church. Preferred rates of repetitive tapping and categorical time production. *Perception & Psychophysics*, 55:443–453, 1994. ISSN 0031-5117. doi: 10.3758/BF03205301. URL <http://dx.doi.org/10.3758/BF03205301>. (Cited on page 34.)
- C Douglas Creelman. Human discrimination of auditory duration. *The Journal of the Acoustical Society of America*, 34:582, 1962. (Cited on page 37.)
- Roger B Dannenberg. An on-line algorithm for real-time accompaniment. In *Proceedings of the 1984 International Computer Music Conference*, pages 193–198. Computer Music Association, 1984. (Cited on page 33.)
- Simon Dixon. An interactive beat tracking and visualisation system. In *Proceedings of the international computer music conference*, pages 215–218, 2001. (Cited on page 38.)
- Carolyn Drake, Mari Riess Jones, and Clarisse Baruch. The development of rhythmic attending in auditory sequences: attunement, referent period, focal attending. *Cognition*, 77(3):251–288, 2000. doi: 10.1016/S0010-0277(00)00106-2. (Cited on page 16.)

Douglas Eck. Finding downbeats with a relaxation oscillator. *Psychological Research*, 66(1):18–25, 2002. (Cited on page 37.)

Douglas Eck. Beat tracking using an autocorrelation phase matrix. In *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, volume 4, pages IV–1313. IEEE, 2007. (Cited on pages 28 and 36.)

W Tecumseh Fitch and Andrew J Rosenfeld. Perception and production of syncopated rhythms. *Music Perception*, 25(1):43–58, 2007. (Cited on page 14.)

R. Fitzhugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophysical journal*, 1(6):445–466, 1961. (Cited on pages 38 and 47.)

James W Haefner. *Modeling biological systems: principles and applications*. Springer, 2005. (Cited on pages 134, 139, and 148.)

Stephen Handel. Using polyrhythms to study rhythm. *Music Perception*, pages 465–484, 1984. (Cited on page 3.)

Stephen Handel and James Oshinsky. The meter of syncopated auditory polyrhythms. *Attention, Perception, & Psychophysics*, 30(1):1–9, 1981. (Cited on pages 6, 22, 69, 73, 74, 76, 80, 82, 83, 87, 89, 90, 105, 113, 123, 155, 156, 190, 191, and 198.)

Alan L. Hodgkin and Andrew F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4):500, 1952. (Cited on pages 37, 38, 46, and 47.)



- Frank C. Hoppensteadt and Eugene M. Izhikevich. Canonical models for bifurcations from equilibrium in weakly connected neural networks. *World Congress on Neural Networks*, 1(2):80–83, 1995. (Cited on page 53.)
- Frank C. Hoppensteadt and Eugene M. Izhikevich. Synaptic organizations and dynamical properties of weakly connected neural oscillators. *Biological Cybernetics*, 75(2):117–127, 1996a. (Cited on pages 44, 48, 50, 51, 54, 55, 56, 80, and 201.)
- Frank C Hoppensteadt and Eugene M Izhikevich. Synaptic organizations and dynamical properties of weakly connected neural oscillators ii. learning phase information. *Biological Cybernetics*, 75(2):129–135, 1996b. (Cited on page 202.)
- Dan H. Hubel, Thomas N. Wiesel, et al. Receptive fields of cells in striate cortex of very young, visually inexperienced kittens. *Journal of neurophysiology*, 26(6):994–1002, 1963. (Cited on page 44.)
- John R. Iversen, Bruno H. Repp, and Aniruddh D. Patel. Top-down control of rhythm perception modulates early auditory responses. *Annals of the New York Academy of Sciences*, 1169(1):58–73, 2009. (Cited on page 106.)
- Ray Jackendoff. Précis of foundations of language: brain, meaning, grammar, evolution. *Behavioral and Brain Sciences*, 26(6):651–665, 2003. (Cited on page 88.)
- Mari Riess Jones. Dynamic pattern structure in music: Recent theory and research. *Perception & Psychophysics*, 41(6):621–634, 1987. (Cited on page 199.)
- Peter B Kahn and Yair Zarmi. *Nonlinear dynamics. Exploration through normal forms*, volume 1. 1998. (Cited on page 54.)

James W. Kalat. *Biological Psychology*. Thomson Wadsworth, 2012. (Cited on page 44.)

Anke Karabanov, Örjan Blom, Lea Forsman, and Fredrik Ullén. The dorsal auditory pathway is involved in performance of both visual and auditory rhythms. *Neuroimage*, 44(2):480–488, 2009. (Cited on pages 199 and 200.)

A.P. Klapuri, A.J. Eronen, and J.T. Astola. Analysis of the meter of acoustic musical signals. *Audio, Speech, and Language Processing, IEEE Transactions on*, 14(1):342–355, 2006. (Cited on pages 28 and 36.)

Ursula Koch and Benedikt Grothe. Hyperpolarization-activated current in the inferior colliculus: distribution and contribution to temporal processing. *Journal of neurophysiology*, 90(6):3679–3687, 2003. (Cited on page 46.)

Gerald Langner. Temporal processing of periodic signals in the auditory system: Neuronal representation of pitch, timbre, and harmonicity. *Z. Audiol*, 46(1):80–21, 2007. (Cited on page 109.)

Edward Large. Musical tonality, neural resonance and hebbian learning. *Mathematics and Computation in Music*, pages 115–125, 2011. (Cited on page 202.)

Edward Large, Philip Fink, and Scott Kelso. Tracking simple and complex sequences. *Psychological Research*, 66(1):3–17, 2002. (Cited on pages 15, 16, and 18.)

Edward W. Large. On synchronizing movements to music. *Human Movement Science*, 19(4):527–566, 2000. (Cited on pages 20, 35, 36, 62, 65, 66, 67, 70, 79, 80, 87, 155, 157, 158, 160, 161, 180, 181, 182, 183, and 190.)

Edward W. Large. Resonating to musical rhythm: theory and experiment. *Psychology of time*, pages 189–232, 2008. (Cited on pages 2, 14, 58, 61, 62, 116, 136, and 193.)

Edward W Large. Neurodynamics of music. In *Music perception*, pages 201–231. Springer, 2010. (Cited on pages 3, 35, 67, and 140.)

Edward W. Large and Mari Riess Jones. The dynamics of attending: How people track time-varying events. *Psychological Review*, 106(1):119–159, 1999. (Cited on pages 35, 39, and 62.)

Edward W. Large and John F. Kolen. Resonance and the perception of musical meter. *Connection Science*, 6(2-3):177–208, 1994. (Cited on pages 28, 33, 34, 35, 62, and 65.)

Edward W Large and John F Kolen. Method and apparatus of analysis of signals from non-stationary processes possessing temporal structure such as music, speech, and other event sequences, May 12 1998. US Patent 5,751,899. (Cited on page 62.)

Edward W. Large and Caroline Palmer. Perceiving temporal regularity in music. *Cognitive Science*, 26(1):1–37, 2002. (Cited on pages 35 and 62.)

Edward W. Large and John S. Snyder. Pulse and meter as neural resonance. *Annals of the New York Academy of Sciences*, 1169(1):46–57, 2009. (Cited on pages v, 2, 35, and 62.)

Edward W. Large, Caroline Palmer, and James B. Pollack. Reduced memory representations for music. *Cognitive Science*, 19(1):53–93, 1995. (Cited on page 65.)

Edward W. Large, Felix V. Almonte, and Marc J. Velasco. A canonical model for gradient frequency neural networks. *Physica D: Nonlinear Phenomena*,

239(12):905–911, 2010. (Cited on pages 3, 35, 38, 41, 47, 53, 55, 56, 63, 80, 88, 97, 133, 139, 150, 151, and 200.)

Jean Laroche. Estimating tempo, swing and beat locations in audio recordings. In *Applications of Signal Processing to Audio and Acoustics, 2001 IEEE Workshop on the*, pages 135–138. IEEE, 2001. (Cited on pages 28 and 31.)

Fred Lerdahl and Ray S Jackendoff. *A generative theory of tonal music*. The MIT Press, 1983. (Cited on pages 15, 16, and 19.)

Justin London. *Hearing in time*. OUP USA, 2004. (Cited on pages 11, 14, 15, 22, 74, and 75.)

Neil P McAngus Todd, Donald J O’Boyle, and Christopher S Lee. A sensory-motor theory of rhythm, time perception and beat induction. *Journal of New Music Research*, 28(1):5–28, 1999. (Cited on pages 28 and 36.)

NP McAngus Todd and Chris Lee. An auditory-motor model of beat induction. In *Proceedings of the International Computer Music Conference*, pages 88–88. International Computer Music Association, 1994. (Cited on page 36.)

J.D. McAuley. *Perception of time as phase: Toward an adaptive-oscillator model of rhythmic pattern processing*. PhD thesis, 1995. (Cited on pages 28 and 33.)

Leonard Meyer and Grosvenor Cooper. The rhythmic structure of music, 1960. (Cited on pages 15 and 19.)

John Albertus Michon. *Timing in temporal tracking*. Institute for Perception RVO-TNO Soesterberg, The Netherlands, 1967. (Cited on page 37.)

- Benjamin O Miller, Don L Scarborough, and Jacqueline A Jones. On the perception of meter. In *Understanding music with AI*, pages 428–447. MIT Press, 1992. (Cited on page 34.)
- Vernon B. Mountcastle. Modality and topographic properties of single neurons of cat's somatic sensory cortex. 1957. (Cited on page 44.)
- J. Nagumo, S. Arimoto, and S. Yoshizawa. An active pulse transmission line simulating nerve axon. *Proceedings of the IRE*, 50(10):2061–2070, 1962. (Cited on pages 38 and 47.)
- RV O'Neill and RH Gardner. Sources of uncertainty in ecological models. *Methodology in Systems Modelling and Simulation, North-Holland, Amsterdam*, pages 447–463, 1979. (Cited on page 125.)
- Alan V Oppenheim and Ronald W Schafer. Digital signal processing, 1975, 1975. (Cited on pages 28 and 35.)
- Richard Parncutt. A perceptual model of pulse salience and metrical accent in musical rhythms. *Music Perception*, pages 409–464, 1994. (Cited on page 39.)
- Aniruddh D. Patel, John R. Iversen, Yanqing Chen, and Bruno H. Repp. The influence of metricality and modality on synchronization with a beat. *Experimental Brain Research*, 163(2):226–238, 2005. (Cited on page 11.)
- Bettina Pollok, Joachim Gross, Katharina Müller, Gisa Aschersleben, and Alfons Schnitzler. The cerebral oscillatory network associated with auditorily paced finger movements. *Neuroimage*, 24(3):646–655, 2005. (Cited on page 88.)
- Karl R Popper. The logic of scientific discovery. *London: Hutchinson*, 1, 1959. (Cited on page 71.)

Dirk J. Povel and Peter Essens. Perception of temporal patterns. *Music Perception*, pages 411–440, 1985. (Cited on pages 28, 29, 30, and 34.)

Jeff Pressing, Jeff Summers, and Jon Magill. Cognitive multiplicity in polyrhythmic pattern performance. *Journal of Experimental Psychology: Human Perception and Performance*, 22(5):1127, 1996. (Cited on pages 22, 74, and 75.)

Pasko Rakic. Local circuit neurons. *Neurosciences Research Program Bulletin*, 13(3):295, 1975. (Cited on page 44.)

Summer K Rankin, Edward W Large, and Philip W Fink. Fractal tempo fluctuation and pulse prediction. *Music Perception*, 26(5):401–413, 2009. (Cited on page 17.)

Bruno Repp. Sensorimotor synchronization: A review of the tapping literature. *Psychonomic Bulletin & Review*, 12(6):969–992, 2005. (Cited on pages 74, 75, 108, and 157.)

Bruno H Repp. Subliminal temporal discrimination revealed in sensorimotor coordination. *Rhythm perception and production*, pages 129–142, 2000. (Cited on page 166.)

Bruno H Repp. Phase correction in sensorimotor synchronization: Nonlinearities in voluntary and involuntary responses to perturbations. *Human Movement Science*, 21(1):1–37, 2002. (Cited on page 18.)

Bruno H Repp. Multiple temporal references in sensorimotor synchronization with metrical auditory sequences. *Psychological Research*, 72(1):79–98, 2008. (Cited on page 75.)

Eric D Scheirer. Tempo and beat analysis of acoustic musical signals. *The Journal of the Acoustical Society of America*, 103:588, 1998. (Cited on pages 28, 36, and 39.)

Eric D Scheirer. *Music-listening systems*. PhD thesis, Massachusetts Institute of Technology, 2000. (Cited on page 38.)

Jarno Seppänen. *Computational models of musical meter recognition*. PhD thesis, 2001. (Cited on pages 28, 31, and 38.)

GM Shepherd. *Local circuit neurons*. MIT Press, 1976. (Cited on page 44.)

CE Sherrick Jr et al. Perceived order in different sense modalities. *Journal of experimental psychology*, 62:423, 1961. (Cited on page 37.)

Joel Snyder and Carol L Krumhansl. Tapping to ragtime: Cues to pulse finding. *Music Perception*, 18(4):455–489, 2000. (Cited on pages 66, 180, 181, 182, and 203.)

Joel S. Snyder and Edward W. Large. Tempo dependence of middle- and long-latency auditory responses: power and phase modulation of the eeg at multiple time-scales. *Clinical Neurophysiology*, 115(8):1885–1895, 2004. (Cited on pages 74 and 75.)

Steven H Strogatz. *Nonlinear dynamics and chaos*. 1994. Reading: Perseus Books, 1994. (Cited on pages 51 and 52.)

Michael Thaut et al. *Rhythm Music and the Brain: Scientific Foundations and Clinical Applications*. Routledge, 2005. (Cited on pages 1 and 36.)

Michael H Thaut, Bin Tian, and Mahmood R Azimi-Sadjadi. Rhythmic finger tapping to cosine-wave modulated metronome sequences: Evidence of subliminal entrainment. *Human Movement Science*, 17(6): 839–863, 1998. (Cited on page 18.)

Petri Toiviainen. An interactive midi accompanist. *Computer Music Journal*, pages 63–75, 1998. (Cited on pages 28, 34, and 39.)

Balth Van der Pol and Jan Van der Mark. Lxxii. the heartbeat considered as a relaxation oscillation, and an electrical model of the heart. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 6(38): 763–775, 1928. (Cited on page 37.)

Leon van Noorden and Dirk Moelants. Resonance in the perception of musical pulse. *Journal of New Music Research*, 28(1):43–66, 1999. (Cited on page 128.)

X.-J. Wang and J. Rinzel. Spindle rhythmicity in the reticularis thalami nucleus: Synchronization among mutually inhibitory neurons. *Neuroscience*, 53(4):899 – 904, 1993. (Cited on page 44.)

Stephen Wiggins. *Introduction to applied nonlinear dynamical systems and chaos*, volume 2. Springer, 2003. (Cited on page 54.)

H. R. Wilson and J. D. Cowan. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Biological Cybernetics*, 13(2):55–80, 1973. (Cited on page 47.)

Hugh R. Wilson and Jack D. Cowan. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, 12(1):1–24, 1972. doi: 10.1016/S0006-3495(72)86068-5. (Cited on pages 7, 37, 47, 50, 54, 81, 126, and 194.)

Arthur T Winfree. *The geometry of biological time*, volume 12. Springer Verlag, 2001. (Cited on page 51.)





## COLOPHON

This document was typeset using the typographical look-and-feel classicthesis developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". classicthesis is available for both L<sup>A</sup>T<sub>E</sub>X and L<sup>y</sup>X:

<http://code.google.com/p/classicthesis/>

Happy users of classicthesis usually send a real postcard to the author, a collection of postcards received so far is featured here:

<http://postcards.miede.de/>

*Final Version* as of April 26, 2014 (classicthesis).



## DECLARATION

---

Put your declaration here.

*Milton Keynes near Bletchley, the city of Codebreakers, August 2013*

---

Vassilis Angelis